The structural and biological properties of Photorhabdus Mns, and CYP6G1-mediated insecticide resistance in Drosophila melanogaster

Jones, Robert

Award date:
2007

Awarding institution:
University of Bath

The Structural and Biological Properties of *Photorhabdus* Mns, and CYP6G1-Mediated Insecticide Resistance in *Drosophila melanogaster*

Robert Jones

A thesis submitted for the degree of Doctor of Philosophy

Department of Biology and Biochemistry
University of Bath

September 2007

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

Photorhabdus colonise the gut of insect-pathogenic nematodes, and are themselves insect pathogens that produce variety of toxicity factors. One protein, Mns, is highly conserved within the genus and constitutes over 30% of the total protein secretion from *Photorhabdus luminescens*, but is shown in this study not to be an insect toxin. Mns associates with extracellular material when *Photorhabdus* grows in colony biofilms, and modifies the attachment of cells to surfaces. The protein is also detected in aggregations of *Caenorhabditis elegans* nematodes induced by *Photorhabdus asymbiotica* supernatant, but appears not to interact with its symbiotic host nematodes, *Heterorhabditis*. Here, we use circular dichroism, dynamic light scattering and differential scanning calorimetry to characterise Mns, and propose that the high propensity of the protein to aggregate *in vitro* may relate to its function *in vivo*.

In this report, cytochrome P450-mediated insecticide resistance in *Drosophila melanogaster* has also been investigated. In Hikone-R flies, resistance to DDT and other insecticides is conferred by overexpression of a single cytochrome P450 gene, *Cyp6g1*. Here, CYP6G1 is purified from *Escherichia coli* as a recombinant protein, and used to prepare anti-CYP6G1 antibodies to show the expression pattern of the protein in insecticide-resistant and -susceptible flies. CYP6G1 is expressed in Hikone-R fly sperm, indicating that the protein has a role in reproduction. Furthermore, a homology model of CYP6G1 is presented, and shows not only how a variety of insecticides can be accommodated by the active site cavity, but also that the enzyme may contribute to hormone titres in the fly. How this relates to life history traits previously identified in Hikone-R flies is discussed.

ACKNOWLEDGEMENTS

I am very grateful for the help I've received. My biggest thanks go to Nick Waterfield, Stefan Bagby, Jean van den Elsen and Richard ffrench-Constant, whose ideas, expertise and time have been invaluable and are much appreciated.

Many people have provided practical assistance and materials during the data collection stages of this report. For this I mostly thank Michelle Hares and Andrea Dowling, Isabella Vlisidou, Guowei Yang and Abhishek Upadhyay, to each of whom I'm very grateful.

I have also benefited from the knowledge and generosity of Stuart Reynolds and Ioannis Eleftherianos, Sam Boundy, Tim Karr, Ben Heath and Kim Steeds. Toby Jenkins and Xavi Munez, Rene Feyereisen, Alan Cooper and Sharon Kelly, Todd Ciche, Phillip Daborn, Ursula Potter, Kevin Balbi and Kay Uppington have all been valuable collaborators and must be thanked. Sandra Barnes, Paul Wilkinson, Rob Jackson, David Clarke and his group, Tim Bearder, Sophie Ainsworth, Matt Amos and Maria Sanchez-Contreras have contributed to this work. Saskia Bakker did the majority of the P450 homology modelling and has been a very helpful colleague. I also have to thank Sally Shuttleworth and Kenneth Holbourn for help in this area. It was a pleasure to work with Caroline McCart, who helped with all of my *Drosophila*-related work and has patiently continued to answer all of my questions.

Lastly I thank my parents and Stewart, Matthew and Zoë, and those who may not have contributed directly, but like many of those listed above, have enriched my time at the University of Bath. Here I include all members past and present of our lab, Ronald Jenner, Peter Millichap, Samantha McCavera, Victoria Eastham, Joanna Clark and Robin Francis, Tamsin Greenway, TeamBath, CrewBath, and the Sports Training Village.

ABBREVIATIONS

AChE	Acetylcholinesterase
amp	Ampicillin
BSA	Bovine serum albumin
CD	Circular dichroism
CFU	Colony forming units
Cif	Cycle inhibiting factor
CNF	Cytotoxic necrotising factor
COSY	Correlation spectroscopy
CR	Congo red
CV	Crystal violet
DDT	Dichloro-diphenyl-trichloroethane
δ-ALA	δ-aminolevulinic acid
DLS	Dynamic light scattering
DNA	Deoxyribonucleic acid
DSC	Differential scanning calorimetry
DTT	Dithiothreitol
EDTA	Ethylenediamine tetraacetic acid
GABAR	γ-aminobutyric acid receptor
GASP	Growth-advantage-in-stationary-phase
GFP	Green fluorescent protein
GST	Glutathione-S-transferase
H_{cal}	Calorimetric enthalpy
H_{VH}	van't Hoff enthalpy
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Hms	Hemin storage
HSQC	Heteronuclear single quantum correlation
HTV	High tension voltage
HWI	Hauptman Woodward Institute
IJ	Infective juvenile
IPTG	Isopropyl-β-D-thiogalactopyranoside
IS	Insertion sequence

JH	Juvenile hormone
kan	Kanamycin
KDR	Knockdown resistance
LB	Luria Bertani Lysogeny Broth
MBP	Maltose binding protein
Mcf	Makes caterpillars floppy
Mns	Makes nematodes sticky
MOE	Molecular Operating Environment
MROD	Methoxy-resorufin O-Demethylation
Mrf	Mannose-resistant fimbriae
MWCO	Molecular weight cut-off
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
NRMSD	Normalised root mean standard deviation
ODM	Octadecylmercaptan
P450	Cytochrome P450
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein databank
PEG	Polyethylene glycol
pI	Isoelectric point
Pnf	*Photorhabdus* necrotising factor
PVC	*Photorhabdus* virulence cassette
Rdl	Resistant to dieldrin
rDNA	Ribosomal DNA, genes coding for ribosomal RNA
RFLP	Restriction fragment length polymorphism
rif	Rifampicin
RT	Room temperature
SAM	Self-assembled monolayer
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEM	Scanning Electron Microscopy
SPR	Surface plasmon resonance
SRS	Substrate recognition site

SVD	Singular value decomposition
T_m	Transition midpoint
TB	Terrific Broth
Tc	Toxin complex
TCA	Trichloroacetic acid
TEM	Transmission electron microscopy
tet	Tetracyclin
TEV	Tobacco etch virus
TOCSY	Total correlated spectroscopy
TPBS	Tween, milk powder PBS
Tris	Tris(hydroxymethyl)aminomethane
TROSY	Transverse relaxation optimised spectroscopy
TT01	*Photorhabdus luminescens* strain TT01
TTSS	Type III secretion system
UV	Ultraviolet
20E	20-hydroxyecdysone

AMINO ACIDS

Alanine	Ala	A	Leucine	Leu	L
Arginine	Arg	R	Lysine	Lys	K
Asparagine	Asn	N	Methionine	Met	M
Aspartic acid	Asp	D	Phenylalanine	Phe	F
Cysteine	Cys	C	Proline	Pro	P
Glutamine	Gln	Q	Serine	Ser	S
Glutamic acid	Glu	E	Threonine	Thr	T
Glycine	Gly	G	Tryptophan	Trp	W
Histidine	His	H	Tyrosine	Tyr	Y
Isoleucine	Ile	I	Valine	Val	V

CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

CHAPTER 7

CHAPTER 1

Introduction to *Photorhabdus*

1 INTRODUCTION

Photorhabdus are bioluminescent, insect-pathogenic Proteobacteria that live in association with soil nematodes. As the nematodes protect and transport the *Photorhabdus*, so the bacteria provide an environment that supports reproduction, supplying nutrients and antimicrobial agents, as well as secreting virulence factors that help to rapidly kill the insect host of the nematode.

1.1 Biochemical and Physiological Characterisation

All species of the *Photorhabdus* genus are motile, nonsporulating rods. They are facultatively anaerobic, with respiratory and fermentative metabolisms, classified as oxidase negative, chemoorganotrophic heterotrophs. Furthermore, all strains are entomopathogenic, and kill the larvae of the wax moth *Galleria mellonella* when injected into the haemocoel. The natural habitat of *Photorhabdus* species is the intestinal lumen of nematodes of the family Heterorhabditae, and the body cavities of insects infected by their vectors (Akhurst & Boemare, 1988). One species, *Photorhabdus asymbiotica*, has also been isolated from human wounds (Gerrard *et al.*, 2003).

1.2 Taxonomy and Phylogeny

The taxonomy of *Photorhabdus* species (formerly members of the genus *Xenorhabdus*) has been relatively well studied, following their placement among the taxon-rich family the Enterobacteriaceae (Thomas & Poinar, 1979). The genus has all the biochemical and morphological characteristics that define the *Enterobacteriaceae*, including production of the enterobacterial common antigen that separates its members from other Gram-negative, facultatively anaerobic rods, the *Vibrioneceae* (Ramia *et al.*, 1982). However, *Photorhabdus* and *Xenorhabdus* species are atypical among their family. Most significantly, they are unable to reduce nitrates to nitrites and fail to produce oxidase. Cell size and colony morphology also differentiate them from other members, as do their symbiotic association with nematodes and pathogenicity towards insects. Other classical taxonomic characteristics, including growth factor requirements and *ab initio* DNA hybridisation experiments, support their classification (Thomas & Poinar, 1979, Grimont *et al.*, 1984).

The most significant event to come from taxonomical studies has been the separation of *Xenorhabdus luminescens* from other members of the genus. Ecologically, the difference is manifest in a difference of nematode vector; all *Xenorhabdus* species are found exclusively with nematodes of the genus *Steinernema* in the family Steinernematidae, with the exception of *X. luminescens*, which associates with *Heterorhabditis* of the Heterorhabditidae (Akhurst & Boemare, 1988). Biochemical and physiological differences also separate *X. luminescens*, including an ability to emit light and a failure to produce catalase, but elevation to genus status came only after DNA-DNA hybridisation experiments revealed no more that 20% DNA reassociation between *X. luminescens* and non-*X. luminescens* strains (Boemare & Akhurst, 1988, Boemare *et al.*, 1993). Strain Hb^T became the type strain for the new genus and for the species *Photorhabdus luminescens*.

More recently, sequencing from 40 strains the 16S rDNA, the gene that encodes the 16S ribosomal subunit, showed a clear phylogenetic separation of *Xenorhabdus* and *Photorhabdus*, with the evolution of *Photorhabdus* apparently preceding the *Xenorhabdus* radiation (Szállás *et al.*, 1997). The phylogeny was reinforced by a polyphasic classification with the power to delineate the genus into three bacterial species: *Photorhabdus luminescens*, *Photorhabdus temperata* and *Photorhabdus asymbiotica,* and a series of subspecies (Fischer-Le Saux *et al.*, 1999). These species have low genomic relatedness (<42% DNA hybridisation), high divergence (thermal stability of DNA duplexes ΔT_m >8.2°C), deep phylogenetic branching and perform differently in phenotypic tests. Furthermore, the data show that the genus is heterogeneous and several of the identified genotypes may constitute more subspecies. Lastly, phylogenetic studies have identified a close correlation between strain grouping and the nematode species from which they were isolated, indicating a coevolution of bacteria and their eukaryotic vector (Liu *et al.*, 1997, Fischer-Le Saux *et al.*, 1999).

1.3 *Photorhabdus* Lifecycle

The *Photorhabdus* lifecycle involves symbiotic and pathogenic activities required for interaction with its two eukaryotic hosts, the nematode vector and the insect it infects. Until recently, no nematode host had been identified for *P. asymbiotica*, which was thought to live free of their association and had been isolated only from human wounds (Gerrard *et al.*, 2003). A strain has now been reported living with an insect-pathogenic species closely related to *Heterorhabditis indica* (Gerrard *et al.*, 2006). The review here will focus on the more established lifecycle of *P. luminescens* and its host *Heterorhabditis bacteriophora* (**Fig. 1.1**).

The cycle begins with an infective juvenile (IJ), a developmental stage of the nematode that is non-feeding, sheathed in a secondary cuticle that protects it from desiccation, and able to move freely in the soil. The gut of the IJ is colonised by up to 300 *Photorhabdus* cells, which are maintained by the nematode but have limited capacity to multiply or spread within the intestine (Ciche & Ensign, 2003). Searching nematodes find their insect hosts by responding to host-associated cues such as CO_2 and faeces, and may be ingested by a feeding insect. Alternatively, contact cues and host acceptance are followed by cuticular penetration of soft body insect larvae, or entry at membranous intersegmental areas of adult insects (Wang & Gaugler, 1998).

Once inside the insect, the nematode sheds its outer cuticle and begins to regurgitate individual cells of *Photorhabdus* (Ciche & Ensign, 2003). The bacteria grow in the haemolymph and midgut, and produce a large number of toxins that rapidly kill the host (Silva *et al.*, 2002). *Photorhabdus* are highly virulent, with just one to two cells required for a 50% lethal dose when injected into larvae of *G. mellonella* and the tobacco hornworm *Manduca sexta* (Milstead, 1979, Bowen & Ensign, 1998). Indeed, sequencing the genome of *P. luminescens* has revealed a trove of factors involved in killing the host insect (Duchaud *et al.*, 2003). These include lipases, two homologs of insect juvenile hormone esterases and a novel toxin designated Mcf, which is encoded on a pathogenicity island (Clarke & Dowds, 1995, Waterfield *et al.*, 2002). The most important appear to be the toxin complex (*tc*) genes, which are active against a wide range of insects from several orders including lepidoptera, coleoptera, diptera and hemiptera, and have been proposed as successors to the Bt toxin in insect-resistant transgenic plants (Bowen & Ensign, 1998, Bowen *et al.*, 1998).

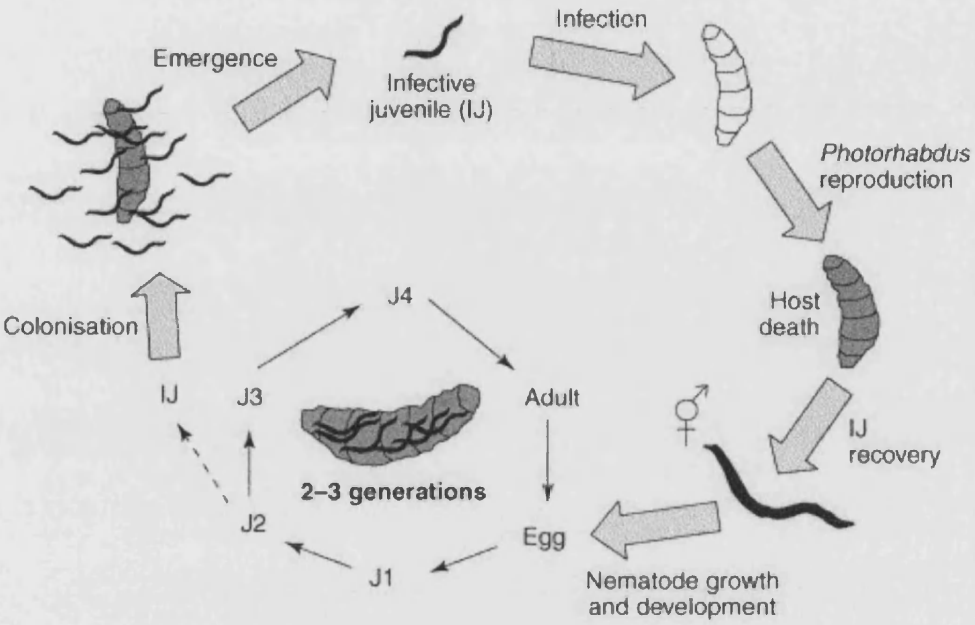


Fig. 1.1 The *Photorhabdus* lifecycle. An infective juvenile nematode infects a soil-living insect larva and regurgitates the *Photorhabdus* carried in its intestine. The bacterial cells reach stationary phase and release toxins into the haemolymph which kill the insect within 24-72h. Meanwhile, the nematode has matured into a hermaphroditic adult that is self-fertile and lays eggs. The juveniles that hatch grow into male and female nematodes that propagate, producing two or three generations that feed on the bacterially-converted biomass. As resources are depleted, IJs develop and are colonised by their specific *Photorhabdus* symbiont. Several thousand IJs emerge from the carcass and search for a new host. (From Joyce *et al.*, 2006).

The toxin complexes effect their damage by creating large cavities in the epithelial midgut of the insect (**Fig. 1.2**). The damage rapidly spreads posteriorly and within 48 hours the lumen of the gut shows total disorganisation, causing the insect to cease feeding and compromising the haemolymph-based immune system. *Photorhabdus* also evade the cellular immune response, apparently by secreting an antiphagocytic factor (Silva *et al.*, 2002).

P. luminescens becomes the dominant flora in the cadaver, and a major food source to nematodes as they bioconvert the insect tissue. Mature, hermaphroditic nematodes develop from IJs, and undergo several rounds of reproduction in the insect (Ciche & Ensign, 2003). In apparent response to the imminent exhaustion of resources, the propagative stage of the life cycle terminates and a new generation of IJs develop which are colonised by *Photorhabdus*. The IJs retain in their intestines only their specific bacterial symbiont strain; the specificity of this retention likely comes from an adhesin-receptor interaction in the intestine, or from surviving the nematode defences that exclude all others (Ciche & Ensign, 2003). Transmission continues when thousands of charged nematodes emerge from the carcass in search of another victim.

P. luminescens is therefore a specialised microbe that lives in a complex lifecycle in which it acts as a nematode symbiont, insect pathogen and nematode food source. Extensive horizontal transfer has occurred and equipped the species with the genes necessary to fill this niche, but how it regulates its insecticidal capabilities and nematode interactions remains unclear.

1.4 Phase Variation

The pathogenicity and symbiotic activities of *Photorhabdus* are temporally separated in the bacterial lifecycle, and Joyce and Clarke (2003) have suggested the traits expressed during these stages are associated with the phenomenon of phenotypic variation. Pathogenic bacteria are able to respond to environmental change at the individual and community level, expressing those genes that will best equip them for invasion, avoidance of host defences and replication. Genetic instability generates phase variants, so preadapting a clonal population to potentially rapid changes in its environment by ensuring that a proportion of cells will be displaying the necessary

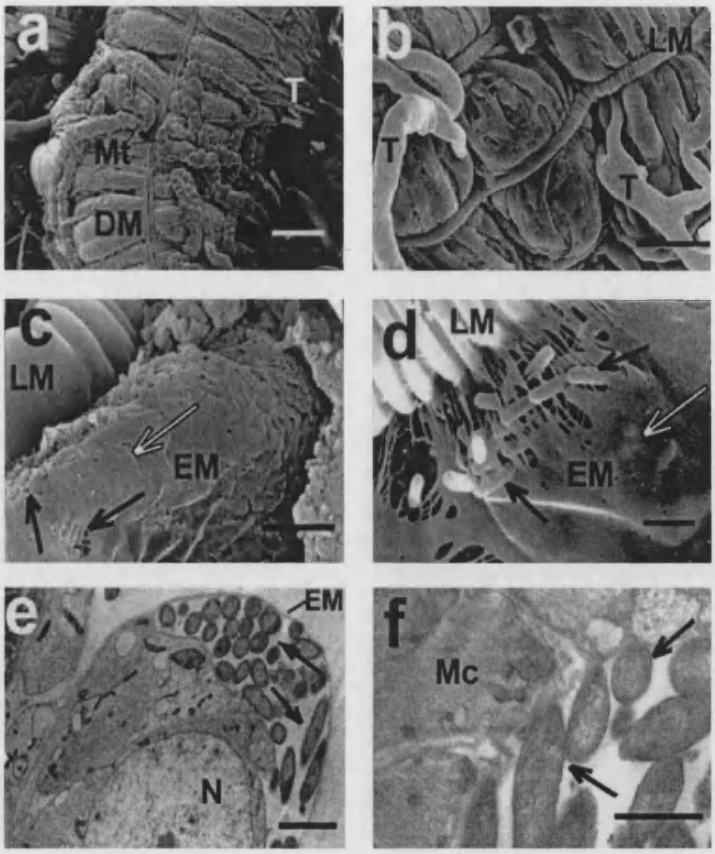


Fig. 1.2 Examination of infected *M. sexta* midgut by electron microscopy. **(A)** An uninfected midgut viewed by SEM shows a folded midgut epithelium bundled together by muscles, including the dorsal muscle (DM). The Malpighian tubules (Mt) lie alongside the midgut and the tracheae (T) branch into the midgut tissues to bring oxygen. **(B)** Longitudinal muscles (LM) hold the folds of gut epithelium together in the uninfected insect. SEM of an infected insect shows bacteria underneath the surrounding extracellular matrix (EM) 48 h post injection. **(C)** and **(D)** Individual bacteria lie alongside each other at high density where the EM has ruptured (black arrows) and the outlines of numerous bacteria can also be seen beneath the EM itself (white arrow). Transmission electron microscopy (TEM) of an infected section of insect midgut shows that the bacteria are underneath the EM in close proximity to the cells of the midgut epithelium itself. **(E)** The gut cells are often detached from the basal lamina, but no obvious cellular rearrangement is observed, despite their close proximity to the invading bacteria **(F)**. Scale bars: A, 500μm; B, 50μm; C, 5μm and D-F, 2μm. (From Silva *et al.*, 2002).

phenotypic traits (Dybvig, 1993). Phase variation is caused by genetic modifications including methylation, slipped-strand mispairing in promoter sites, random mutations and genomic rearrangements, such as the transposition of insertion sequence (IS) elements. The process is reversible and occurs at high frequency (reviewed by Henderson *et al.*, 1999). Typical variations include the production of flagella and changes to surface antigens, and are met in the host by similar strategies of genetic rearrangements that allow rapid recognition of the pathogen (Brunham *et al.*, 1993).

1.4.1 *Photorhabdus* Phase Variants

In *Photorhabdus*, primary variants (phase I) are isolated from field-collected infective-stage nematodes and insect carcasses but are unstable *in vitro* and produce the secondary (phase II) variant after prolonged incubation (Akhurst, 1980). Since secondary variant *Photorhabdus* cells do not revert to the primary form, the process has been described as phenotypic variation, rather than phase variation. The phenotypic variants of *Photorhabdus* are universally distinguished by colony morphology, the absorption of neutral red dye from agar, and the production of antibiotic molecules (Boemare & Akhurst 1988). The variants also differ in several other characteristics including motility, degree of pigmentation and bioluminescence, lecithinase, protease and lipase secretion and crystal protein production (reviewed by Forst 1997).

The characteristics of primary variant cells are only expressed when cultures have reached the post-exponential growth phase. Indeed, a proteomic analysis of primary- and secondary-form cells has shown identical protein profiles during exponential growth (Turlin *et al.*, 2006). In the insect, exponential growth is associated with pathogenicity, while entry to stationary phase coincides with the establishment of a monoxenic culture and the progression of the nematode lifecycle. Primary cells are therefore associated with the symbiotic activities of *Photorhabdus* and primary-specific characters have been termed symbiosis factors (Joyce & Clarke, 2003). Secondary-variant cells, characterised by an absence of symbiotic factors, have been isolated from the haemocoel of insects but do not appear to be retained by IJ nematodes and are less able to support nematode growth and reproduction (Akhurst 1980, Han & Ehlers, 2001). Interestingly, the phase variants are equally pathogenic to insects (Bintrim & Ensign, 1998).

What role the secondary variant has in the ecology of *Photorhabdus* is not understood. However, isolation of phase II cells from conditions of starvation, low osmolarity and anoxia incurred after prolonged incubation suggests a possible adaptation to survival outside of the nematode host. Consistent with this hypothesis, secondary cells show elevated cellular metabolism and biomass construction, and more rapidly recover from starvation than do primary variant cells (Smigielski *et al.*, 1994, Turlin *et al.*, 2006). It has therefore been argued that, given the appropriate nutritional factors, phase II cells may revert to the primary form and re-establish their symbiotic activities.

1.4.2 Causes of Phenotypic Variation

Attempts to achieve the reversion of secondary cells to the primary variant may have failed because of the nature of the mechanisms responsible for the initial phenotypic switch, which may prove to be irreversible; unlike in classical phase variation, the phenotypic variation of *Photorhabdus* does not appear to involve unstable DNA rearrangements. Boemare *et al.*, (1993) showed 100% DNA relatedness between the two variants, and RFLP analyses have found no evidence of inversions or deletions (Joyce & Clarke, 2003). These findings also make it unlikely that the secondary variant derives from accumulated mutations akin to those observed in growth-advantage-in-stationary-phase (GASP) mutants (Finkel, 2006).

Recently, regulatory proteins have been implicated in the control of phenotypic variation of *Photorhabdus*. The global regulator H-NS shows growth-phase dependent expression and appears to control genes involved in the response to osmotic and oxidative stress (Turlin *et al.*, 2006). Similarly, *hexA* and *ner* are up-regulated in secondary cells and repress primary characteristics through a complex regulatory cascade (O'Neill *et al.*, 2002, Joyce & Clarke, 2003). Expression of *hexA* also appears to be required for the pathogen-symbiosis transition.

1.4.3 Proteomics of Phenotypic Variation

A proteomic analysis of phenotypic variants of *P. luminescens* has revealed that secondary-phase cells have reduced protein synthesis and may be defective in protein secretion. The most highly secreted protein of primary cells, comprising greater than 30% of the protein complement in the extracellular fraction, was identified as Plu1537, an unknown protein that shows weak similarity to the Cry34A toxin of *Bacillus thuringiensis* (Turlin *et al.*, 2006). Plu1537, referred to in the present study as Mns, is secreted with 88 other proteins in the post-exponential phase of growth. The biological and structural properties of Mns are a focus of this report and are covered in subsequent chapters.

1.5 Virulence Mechanisms of *Photorhabdus*

The potential use of *Photorhabdus* in the control of pest insects makes study of their toxins and mode of action of considerable practical importance. The isolation of strains from human clinical specimens has also encouraged study of the evolution of human pathogens and elevated the need to understand their virulence mechanisms. The genome sequence of *P. luminescens* subspecies *laumondii* strain TT01 contains 4,839 protein-coding genes and has more predicted toxin genes than any other fully sequenced bacterial genome (Duchaud *et al.*, 2003). It is vital to the lifecycle of *Photorhabdus* that the insect host is killed, allowing transmission of the nematode vector, and this may explain the apparent functional redundancy of toxins to 'overkill' the insect. Furthermore, the toxins likely have different target sites, or show different levels of activity in different types of insect hosts. Lastly, toxins are required to keep the insect free of microbial competitors.

1.5.1 Toxin Complexes

The toxin complexes are high molecular weight proteins with insecticidal effects specific to the midgut epithelium. The complexes are delivered directly into the insect haemocoel once the bacteria have been released from their nematode host. Unexpectedly, the proteins are also orally toxic and can interact with the lumen side of the insect midgut (Blackburn *et al.*, 1998). Furthermore, disruption of the genes encoding the toxin complexes in *Photorhabdus* abolishes their ability to kill *M. sexta* larvae by oral administration; the *tc* genes are necessary and sufficient for oral toxicity of *Photorhabdus* (Bowen *et al.*, 1998).

In *P. luminescens* strain W14, four loci were originally identified that encode the Tc toxins (*tca*, *tcb*, *tcc* and *tcd*), but there is conservation within these loci and only three basic genetic elements are evident: *tcaAB/tcdA*-like genes, *tcaC/tcdB*-like genes and *tccC*-like genes (Waterfield *et al.*, 2001). The activity of the proteins they encode, and the composition and stoichiometry of the native complexes they form, is not fully understood. However, Waterfield *et al.* (2005) have described a model in which TcdA interacts with the insect midgut epithelium and shows toxic activity that is enhanced by the coexpression of TcdB and TccC.

Similar Tc loci have been identified in other bacteria, including *Xenorhabdus nematophilus*, *Serratia entomophila*, *Yersinia pestis* and *Yersinia pseudotuberculosis* (Morgan *et al.*, 2001, Hurst *et al.*, 2000, Parkhill *et al.*, 2001) as well as *Pseudomonas syringae* pv syringae and *Pseudomonas fluorescens* and may be used in defence against insects they may encounter in their lifecycles (Morgan *et al.*, 2001, Hurst *et al.*, 2000, Parkhill *et al.*, 2001, ffrench-Constant & Waterfield, 2006). Interestingly, the potentiator genes *tcdB* and *tccC* are found in pairwise loci in *Y. pestis* and *S. entomophila*, raising the formal possibility that the potentiators are themselves toxic, and have a role independent of TcdA.

During *Photorhabdus* infection, toxin complexes are produced within the gut niche and are displayed on the extracellular surface of the bacteria. Proliferation of the pathogen ensures that high local concentrations accumulate, creating large cavities and the release of cellular debris into the lumen. The histopathological effects are similar to those caused by the δ-endotoxins of *B. thuringiensis* (Silva *et al.*, 2002, Blackburn *et al.*, 1998).

The toxin complexes are therefore of potential use in crop protection because they are active when administered orally and may be used to delay the evolution of resistance to the *B. thuringiensis* toxins which are widely deployed in insect-resistant transgenic plants (ffrench-Constant & Waterfield, 2006). Moreover, the *tc* genes show activity against a broader range of insect hosts and are of comparable potency; in the Colorado potato beetle, *Leptinotarsa decemlineata*, and sweet potato whitefly, *Bemisia tabaci*, Toxin complex a (Tca) causes mortality at concentrations below 1 part per million (Blackburn *et al.*, 2005).

1.5.2 Makes Caterpillar Floppy Toxins

Screening of a *P. luminescens* cosmid library identified a pathogenicity island that enables persistence of recombinant *Escherichia coli* within *M. sexta*, and causes a rapid loss of body turgor. This 'floppy' phenotype is associated with the *mcf* (makes caterpillars floppy) gene, which is conserved in the genus and is surrounded on the *Photorhabdus* chromosome by other virulence factors (Daborn *et al.*, 2002a)

Recombinant *mcf* causes apoptosis of insect haemocytes, macrophage-like cells that patrol the haemocoel, inhibiting the encapsulation and nodule formation that allow for melanisation and death of wild-type *E. coli*. In addition to disintegration of actin in the haemocyte cytoskeleton, Mcf causes an apoptotic effect in mammalian cells *in vitro* and is characterised by membrane blebbing and nuclear fragmentation. These effects are initiated within six hours, and can all be inhibited by the caspase inhibitor zVAD-fmk (Dowling *et al.*, 2004, Daborn *et al.*, 2002a).

The loss of insect body turgor is also associated with the Mcf-induced apoptosis. The protein interacts with the haemocoel side of the insect midgut, causing elongation and blebbing of both columnar and goblet cells. These cells are involved with osmoregulatory activities, and their destruction is consistent with the establishment of the floppy phenotype (Daborn *et al.*, 2002a). How Mcf induces apoptosis remains unclear, but likely involves the BH3-like domain contained in the N-terminus of the protein. BH3-only proteins interact with factors in the outer mitochondrial membrane and promote programmed cell death. The use of a BH3-like domain by a bacterial pathogen to induce the same response would be unprecedented (Dowling *et al.*, 2004).

Recently, a second Mcf factor has been identified that also kills insects. The two proteins share a central domain with the *Clostridium difficile* Toxin B but have variable termini (Waterfield *et al.*, 2003). The Mcf2 N-terminus shows similarity to a type III effector protein HrmA from *P. syringae*, and localises the protein to the nucleus when expressed as a fusion protein in mammalian cells. The fusion protein also causes destruction of the transfected cells. Whether this represents functional redundancy, or if the two proteins have differing target sites or target species, or different modes of action, remains to be determined.

1.5.3 *Photorhabdus* Virulence Cassette

A further class of insecticidal toxins are encoded in the *Photorhabdus* virulence cassettes (PVCs). These are functional homologs of a prophage-like locus of *S. entomophila* that causes amber disease in *Costelytra zealandica* larvae by inducing an antifeeding effect and clearance of the midgut. The locus encodes phage-like tail and base-plate proteins, and a putative antifeeding effector toxin likely acquired through horizontal gene transfer (Hurst *et al.*, 2004). Correspondingly in *Photorhabdus*, the PVC unit contains 13-15 genes with high levels of similarity to components of phage, and a number of putative effector genes at the 3' end of the prophage cluster. Each strain examined carries several copies of PVC loci, each with different effector genes (Yang *et al.*, 2006). In *P. asymbiotica*, the predicted amino acid sequences of PVC effectors include homologs of *Photorhabdus* proteins Tcc and LopT and the *E. coli* proteins cytotoxic necrotising factor 1 (CNF-1) and cycle inhibiting factor (Cif).

Expression of PVC elements in *E. coli* produces large needle-like structures protected by an outer sheath up to 800nm long (Yang *et al.*, 2006). These show structural similarity to a type of bacteriocin, the R-type pyocins, and are hypothesised to act by contraction of the outer sheath and injection of the exposed needle through the outer membrane of a bacterial cell (Michel-Briand & Baysee, 2002). In the case of the PVC elements, there is no apparent bactericidal activity. However, the PVC structures are injectably toxic to insects (Yang *et al.*, 2006). The CNF-1 homolog, *Photorhabudus* necrotising factor (Pnf) encoded on *P. asymbiotica* PVCpnf, causes condensation of actin filaments in insect haemocytes, and causes nuclear disintegration and disruption of the cytoskeleton when expressed in mammalian NIH-3T3 cells.

CNF modulates Rho family GTPases, which cycle between inactive and active states and act as molecular switches in many signal transduction pathways. When modulated by CNF, Rho GTPases become locked in the active state and are unable to regulate cytoskeletal organisation, inhibiting bacterial phagocytosis and cell-cycle progression (Nougayrède *et al.*, 2005). Cif is also an activator of Rho GTPases and reorganises the cytoskeleton, inducing nuclear enlargement and endoreduplication. The PVC-associated Cif homolog of *Photorhabdus* has not been studied.

Lastly, the *P. luminescens* TT01 genome encodes two genes encoding proteins similar to insect juvenile hormone esterases. By inactivating juvenile hormone, the hormone that maintains the juvenile state, *plu4092* and *plu4436* of *P. luminescens* may kill insect hosts by modifying their development (Duchaud *et al.*, 2003).

1.5.4 Antimicrobial Toxins

In addition to the factors necessary to kill an insect, *Photorhabdus* produces a range of antibiotics that prevent colonisation by microbial competitors. Genes similar to those required by *P. syringae* to synthesise syringomycin and by *Burkholderia glumae* to synthesise the antibiotic toxoflavin have been found in the *P. luminescens* genome, and provide protection against competition by Gram-positive bacteria and fungi which may be present in the soil (Duchaud *et al.*, 2003). Numerous genes predicted to be involved in the transport of these antibiotic factors have also been identified.

Photorhabdus must also exclude closely related bacteria, both in an insect cadaver and within its carrier nematode. Lumicins are novel DNase and RNase bacteriocins produced by *Photorhabdus* that can kill conspecifics but are not toxic to the host bacteria. As is the case of the colicins of *E. coli* and pyocins of *Pseudomonas aeruginosa*, the killing genes of lumicins are genetically linked to immunity genes that deactivate the toxins. Thus, the host cell and clonal members of the same strain are resistant but other strains are potentially susceptible and may be eliminated from the cadaver (Sharma *et al.*, 2002). The maintenance of a single-strain infection is important to the establishment of a monoculture within the nematode, but will also enhance competitive infectivity. Mixed infections, in which two strains can kill each other, are less virulent than single-strain infections (Massey *et al.*, 2004).

1.5.5 Proteolytic Activity

Four secreted proteins with proteolytic activity have been identified in *Photorhabdus* species, the most dominant of which is PrtA, an RTX-family metalloprotease (Bowen *et al.*, 2003). During infection, proteolytic activity is observed 20-40h post-inoculation, initiating well before the onset of mortality. However, PrtA alone is not toxic and there is considerable variation in both the composition and intensity of secretion of proteases between different strains; the most highly pathogenic lack some of the proteolytic activities exhibited by less pathogenic strains (Marokhazi *et al.*, 2004b). There is also no indication that these proteases, like AprA of *Pseudomonas entomophila*, are required to protect the pathogen from the insect immune system by inhibiting antibacterial peptides (Liehl *et al.*, 2006). It therefore appears unlikely that secretion of proteases is directly important to pathogenicity. In *Photorhabdus*, the proteolytic activities of these enzymes have been partially characterised but their natural substrates have not been identified. Their probable function is in the degradation of proteins from the insect host for nutrition; PrtA in the haemocoel specifically adheres to the basal lamina, and has been postulated to break down this membrane to facilitate tissue invasion and bioconversion of the cadaver (Silva *et al.*, 2002). Furthermore, Php-B from *P. luminescens* shows collagen-peptidase activity, and may contribute to bioconversion by hydrolysing fragments derived from collagen that surrounds organs of the insect (Marokhazi *et al.*, 2004a).

1.5.6 Avoidance of Host Defences

Once inside the haemocoel, *Photorhabdus* are susceptible to detection by the host immune system. In *M. sexta*, the pathogen is recognised by immune recognition proteins and a response is mounted that includes production of antibacterial peptides such as cecropin from the fat body, and induction of the prophenoloxidase cascade (Eleftherianos *et al.*, 2006). The activation of prophenoloxidase to active phenoloxidase normally results in the deposition of melanin around invading bacteria by haemocytes. As the haemocytes engulf the bacteria, they aggregate with necrotic cells and coagulated haemolymph, forming nodules that isolate the pathogen.

These antimicrobial defences are ultimately unsuccessful against *Photorhabdus* invasion. However, they do slow down the rate at which the bacteria proliferate and extend the survival time of the insect (Eleftherianos *et al.*, 2006). How the pathogen survives the immune reaction of its insect host is not fully understood, but likely involves activity toward the antimicrobial enzyme phenoloxidase, and the ability to form melanotic nodules. These defence activities are functionally linked, and both are inhibited by a phenylpropanoid hydroxystilbene compound, ST. The compound is first secreted early in natural infections and is important to the virulence of *Photorhabdus* against *M. sexta* (Eleftherianos *et al.*, 2007). Interestingly, ST was initially associated with exclusion of microbial competitors from the cadaver because it shows broad-spectrum antimicrobial activity (Williams *et al.*, 2005).

Photorhabdus are also able to avoid phagocytosis, and dramatically decrease the viability of insect haemocytes (Au *et al.*, 2003). Phagocytosis of bacteria is driven by rearrangements of the actin cytoskeleton, a process regulated by small GTPases of the Rho family. During insect infection, modifications to Rho GTPases inhibit phagocytosis, and are mediated by PVC proteins and LopT, an effector secreted by the type III secretion system (TTSS) of *Photorhabdus*. LopT contains the consensus Cys/His/Asp residues possessed by a novel family of bacterial cysteine proteases, and is homologous to the *Yersinia* protease YopT. The two toxins likely cause the same modifications to the C-terminal domain of their RhoA targets (Brugirard-Ricaud *et al.*, 2005). LopT is secreted during the infection process in a site-specific manner and is delivered into insect cells through the TTSS. Together, the antiphagocitic activity and inhibition of the phenoloxidase system expressed by *Photorhabdus* ensure the insect immune response is sufficiently depressed while the bacteria proliferate and kill their host by deployment of the wide range of insecticidal toxins.

1.6 Bioluminescence

A feature unique to *Photorhabdus* species among terrestrial bacteria is the ability to emit light. This property is not shared with their closest relatives, the *Xenorhabdus*, despite their comparable ecologies. The luciferase genes of *P. luminescens* are very similar in sequence and arrangement to those of the marine bacterium *Vibrio harveyi*, and likely evolved from a common ancestor (Szittner & Meighen, 1990). The benefit that *Photorhabdus* gains from the ability to emit light is unknown, but its occurrence in stationary phase is indicative of a symbiotic function.

1.7 *Photorhabdus*-Nematode Interaction

Photorhabdus reside in the anterior region of the IJ intestine, immediately posterior to the basal bulb, and appear to be actively retained in this location by the nematode (Ciche & Ensign, 2003). In the intestines of non-IJ stages, *Photorhabdus* are digested and appear to provide favourable nutritional conditions for nematode growth and development. Several *Photorhabdus* genes have been identified that are important to this process, including those encoding crystalline inclusion proteins, which may be a nutrient resource (Bintrim & Ensign, 1998), and a *hexA* homolog, which has been implicated in the control of phase characteristics and is of significance because of the inability of secondary phase cells to support nematodes (Joyce & Clarke, 2003). Additionally, mutation of a phosphopantetheinyl transferase homolog, *ngrA*, reduces nematode growth and reproduction (Ciche *et al.*, 2001). The *ngrA* product is involved in the synthesis of antibiotics and iron-chelating siderophores, but its role in nematode development may result from its synthesis of a hormone or signal regulator.

Ciche *et al.*, (2001) noted that the *ngrA* gene is located downstream of putative fimbrial biogenesis factors, which may explain the specificity of the interaction between *Photorhabdus* and their nematode host; IJ juveniles are much better able to carry their own bacterial isolates than those of a different symbiont group (Han *et al.*, 1991, Fischer-Le Saux, 1999). Fimbriae have an important role in the initial attachment of many bacterial pathogens to eukaryotic cells, and are therefore strong candidates for establishing the specific interactions with the nematode host (Edwards & Puente, 1998). Following attachment, signalling may induce changes in gene expression and development in both associates.

The genome of *P. luminescens* reveals further putative adhesion factors, including a protein similar to the *P. aeruginosa* lectin PA-I, which recognises host-cell carbohydrates, and a protein similar to a *Dictyostelium discoideum* calcium-dependent cadherin (Duchaud *et al.*, 2003). A microarray analysis of strains from two host nematode species has identified another candidate region, the *lsr* operon, which is hypothesised to terminate cell-cell (quorum) signalling, and may be important to physiological switches occurring once inside the nematode (Gaudriault *et al.*, 2006). A connection between quorum signalling and fimbrial-mediated adhesion has recently been identified in *Serratia marcescens* (Labbate *et al.*, 2007).

1.8 Other Nematode Interactions

Another example of a specific bacterial interaction with a nematode host has been described for a new species, *Microbacterium nematophilum* (Hodgkin *et al.*, 2000). These bacteria are a potential food source for *Caenorhabditis elegans*, and use an unusual strategy for avoiding predation; by attaching at the post-anal cuticle of the nematode, *M. nematophilum* induce a swelling that increases gut distension and restricts feeding rates. The bacteria are also protected from grazing and have an aid to dispersal. Several genes affect the sensitivity of *C. elegans* to adherence by *M. nematophilum*, including surface antigenicity *srf* factors (Gravato-Nobre *et al.*, 2005). The best characterised of these, *srf-3*, encodes a nucleotide sugar transporter that glycosylates components of the cuticle that coats the nematode. Mutation of *srf-3* prevents adherence by *M. nematophilum*, and it is hypothesised that these bacteria are able to specifically recognise or adhere to the glycosylated components to initiate infection (Höflich *et al.*, 2004).

Alteration of the *C. elegans* surface antigenicity also prevents adherence by *Y. pseudotuberculosis* (Höflich *et al.*, 2004). Adherence of these bacteria and *Y. pestis*, causative agent of bubonic plague, is characterised by the formation of agglomerations around the head and tail of the nematode (Darby *et al.*, 2002). These aggregations impair nematode feeding and development, and require expression of the hemin storage (*hms*) operon, which is also involved in the production of a matrix in the proventriculus of infected fleas. This biofilm-like matrix blocks the insect foregut, causing starvation and maximising transmission of the bacteria as the flea attempts to feed (Hinnebusch *et al.*, 1996).

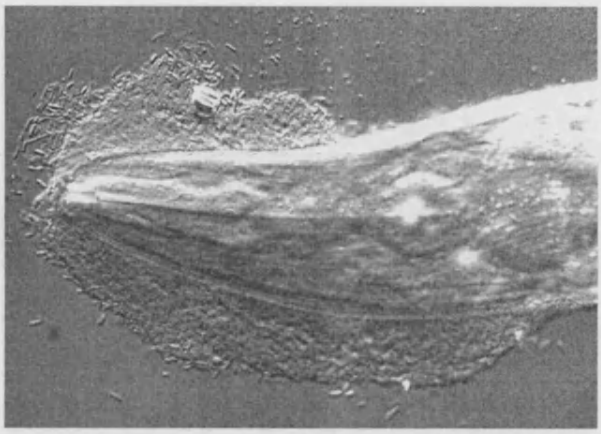
When grown on abiotic surfaces, *Yersinia* species do not readily form biofilms. Joshua *et al.* (2003) found a subset of strains were able to colonise polystyrene, but there was no correlation between these strains and those able to form biofilms on a biotic surface, the head of *C. elegans*. Mucoid *P. aeruginosa* and uropathogenic *E. coli* both fail to produce biofilms on *C. elegans*, further indicating that the process is mediated by specific interaction between bacterial and host recognition molecules (Tan & Darby, 2004). However, the *Yersinia* aggregation appears to be formed from pre-secreted material, and becomes thicker as the moving nematode accumulates additional material; there may be little requirement for signalling between these two organisms or for the trapped cells to contribute to formation of the biofilm.

Biofilm formations on the surface of *C. elegans* have not been described for *Photorhabdus*, but the possibility exists that these bacteria may interact in such a way with other species of nematode. *Xenorhabdus nematophila*, symbiont of nematodes of the family *Steinermatidae*, has been shown to form biofilms similar to those of *Yersinia* around the head and tail of *C. elegans* (**Fig.1.3**) (Couillault & Ewbank, 2002, Sicard *et al.*, 2007). These biofilms on nematodes not normally associated with the bacteria may be involved in environmental persistence and provide an anti-grazing mechanism that protects exposed bacteria from nematodes that will not support their lifecycle

1.9 Biofilms

Biofilm formation at the head of a nematode may result from initial attachment of secreted material, and grow progressively thicker as the nematode moves through and collects further secretions and bacterial cells. On other surfaces however, biofilms can develop into highly structured communities, in which microcolonies are established and the resident bacteria show functional heterogeneity.

(A)



(B)

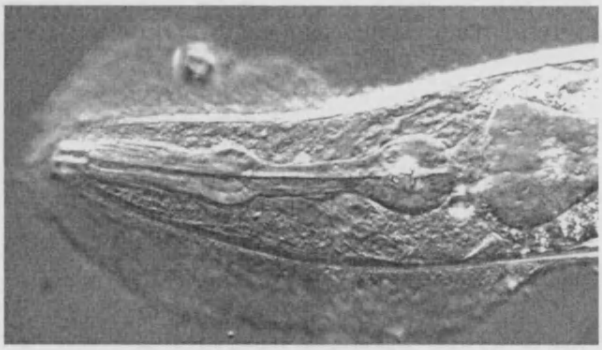


Fig. 1.3 Adhesion of *X. nematophila* to the head of *C. elegans*. Photomicrographs of *C. elegans* that were incubated in the presence of *X. nematophilus* for 2 days and thoroughly washed before being mounted for observation. The images were taken at two different focal planes, to reveal the adherent rod-like bacteria **(A)** and the two-lobed structure of the nematode pharynx **(B)**. *C. elegans* is not the nematode symbiont of *X. nematophilus*, and the formation of this biofilm may be a mechanism to prevent herbivory. (Adapted from Couillault & Ewbank, 2002).

Biofilms enhance the survival of bacterial communities living in hostile environments. The close proximity of cells facilitates intercellular communication and metabolic cooperation, and may increase genetic diversity through horizontal gene transfer (Parsek & Greenberg, 2000, Hausner & Wuertz, 1999). Furthermore, the extracellular polymeric material that surrounds a biofilm provides the bacteria with protection against unfavourable conditions, such as desiccation, chemical biocides and the immune response of a host. Indeed, the resistance of bacterial biofilms to antibiotics, antibodies and phagocytosis allows many human infections to persist (Costerton *et al.*, 1999). The most intensively studied biofilms are those formed by *Pseudomonas aeruginosa*. Development of these biofilms involves firstly reaching the surface, followed by stages of attachment, microcolony formation, and differentiation into mature structures encased in exopolysaccharide. At the air-surface interface, *P. aeruginosa* forms a dense mat called a pellicle.

The initial interactions between bacteria and the substratum are governed by physicochemical factors, such as Brownian motion, van der Waals forces and hydrogen bonding (Busscher & van der Mei, 1997). In *P. aeruginosa*, subsequent events involve two surface associated appendages; a single, polar flagellum and type IV pili (O'Toole & Kolter, 1998). Pili are required for a form of surface-associated, flagella-independent bacterial movement called twitching, which involves extension and retraction of the pili to propel the cell across a thin layer of liquid on a solid surface (Bradley, 1980). The activity also conceivably pulls cells towards a surface (Pasmore & Costerton, 2003). It is hypothesised that this twitching activity is required for *P. aeruginosa* cells to form microcolonies, by bringing cells together into multicell aggregates following homoserine lactone-mediated cell-to-cell signalling (O'Toole & Kolter, 1998).

Signalling between cells is also important in later stages of biofilm development and differentiation (Davies *et al.*, 1998). A *lasI* mutant that doesn't produce homoserine lactone shows normal attachment to an abiotic surface but retains a glycocalyx, the extracellular polymeric material that coats the bacteria, distribution similar to that of planktonic cells. It appears that the accumulation of the signal triggers the differentiation from attached planktonic cells into biofilm bacteria. Non-motile mutants of *P. aeruginosa* that produce a flagellum truncated at the hook, the adaptor

between the rotor and the filament, are unable to found a stable interaction with an abiotic surface (O'Toole & Kolter, 1998). Flagella therefore seem necessary in establishing the first stages of biofilm formation, either adhering through an attractive moiety at the tip of the filament, or by providing necessary motility to overcome repulsive forces between the cell and solid surface. Once attached, pili promote the aggregation of the cells that is required for maturation.

E. coli are peritrichously flagellated, and show a different mechanism in early biofilm formation. Pratt and Kolter (1998) created individual mutant strains defective in the three distinct functions of flagella; motility, adhesion and the chemotactic response, and showed that chemotaxis is dispensable for normal biofilm development. Further, paralysed and non-flagellated cells were able to adhere, albeit poorly, to an abiotic surface. The proposed model for *E. coli* attachment thus begins with cells being brought into contact with the surface through flagella-mediated motility. Type I pili are necessary for the subsequent attachment, and motility at the surface allows the biofilm to expand (Pratt & Kolter, 1998).

Pili and flagella have also been implicated in the adherence of bacteria to eukaryotic cell surfaces (Doig *et al.*, 1988, O'Toole & Kolter, 1998). This suggests a degree of overlap of functions required for both biofilm formation and pathogenesis, and may have implications for *Photorhabdus*. *P. luminescens* has at least eleven clusters of fimbrial genes, including mannose-resistant pili (*mrf*), homologous to the MR/P fimbriae characterised in *Proteus mirabilis*, and type IV pili (Duchaud *et al.*, 2003).

In *P. temperata*, mannose-resistant pili are produced in an insect host in mid-exponential growth, and may play a role late in the infection process (Meslet-Cladiere *et al.*, 2004). Furthermore, the *mrf* pilus operon is flanked by a repressor of flagella synthesis, which is expressed early in infection and suggests that flagella are not involved in interaction with the host. Although the role of these appendages remains unclear, it is of interest here that mannose-resistant pili, which are related to the type I pili of *E. coli*, have been implicated in interbacterial interaction and pellicle formation (Li *et al.*, 2002).

Lastly, two other factors implicated in the development of biofilms are surfactants and the extracellular polymeric substance that encloses the mature structure. Surfactants are surface-active compounds with amphipathic properties that interact with interfaces. A surfactant layer attached to the outside of a bacterial cell allows interaction with either hydrophobic or hydrophilic surfaces, depending on the orientation of the compound. Alternatively, the surfactant may be excreted, so covering an interface with a conditioning film that affects the adhesion or de-adhesion of the bacteria (Neu, 1996). In biofilm formation, surfactants facilitate bacterial movement and are necessary for the development of mushroom-like structures (Pamp & Tolker-Nielsen, 2007). Surfactants also preserve the channels that separate these mushroom formations, and are involved in the detachment of cells from the biofilm, putatively in response to nutrient depletion or the accumulation of a quorum signal (Boles *et al.*, 2005, Hunt *et al.*, 2004).

The extracellular polymeric materials that form a matrix around microbial biofilms are extensions of the glycocalyx, and are akin to the collagen and polysaccharide-based connective tissue that provides support and facilitates attachment in animals. In *P. aeruginosa*, the matrix is composed primarily of carbohydrates, in particular glucose-rich polymers, as well as water, proteins and ions (Friedman & Kolter, 2004). The pellicles of other species, including *E. coli* and *S. typhimurium*, are rich in cellulose, but are not fully characterised (Zogaj *et al.*, 2001). In *Photorhabdus*, pellicle-forming mutants have been isolated and the genes responsible await identification (M.R. Amos, personal communication). The possibility exists that several of the factors outlined above may be involved in biofilm formation. Indeed, the factors responsible may prove to be essential also for nematode interaction, or for insect colonisation during infection.

1.10 Human Infection

The first identification of *Photorhabdus* in isolates from human clinical specimens was reported in 1989. Farmer *et al.* disclosed case reports of patients with wounds infected with *P. asymbiotica* (formerly *Xenorhabdus luminescens* DNA Hybridisation Group 5) but the source of the infections was unknown. Fourteen cases have now been documented in the literature and until recently *P. asymbiotica* had only been isolated from humans (Weissfeld *et al.*, 2005). No nematode associate was known for this species.

The majority of cases of human acquisition of *P. asymbiotica* have occurred during outdoor activity. To search for a nematode host, Gerrard *et al.*, (2006) baited with insect larvae soil taken from the site at which a patient had become infected. Nematodes of the genus *Heterorhabditis* emerged from the insect cadaver 14 days later, and were found to harbour *P. asymbiotica.* The bacterial and nematode strains were denoted Kingscliff strains, and growth and reproduction of these nematodes can only be supported by this strain of *Photorhabdus.* It is unclear whether Kingscliff nematodes are able to penetrate undamaged human skin, or whether a wound is required for entry of the nematode and pathogen. Indeed, a number of cases have been associated with possible spider bites (Gerrard *et al.*, 2003).

The infection caused by *P. asymbiotica* is presented as an invasive, and sometimes multifocal, soft tissue abscess with possible haemotogenous spread and a tendency to relapse. Documented infections have healed after prolonged antimicrobial therapy (Weissfeld *et al.*, 2005). *P. asymbiotica* is therefore seen as a potentially important human pathogen. Its emergence from a genus of invertebrate, rather than vertebrate, pathogens is unusual but is not unique, and is akin to the ecological adaptation of *Bacillus anthracis*, the causative agent of anthrax. A close relative of *B. anthracis* is a virulent insect pathogen *B. thuringiensis* (Ivanova, 2003).

There is no information to suggest that other species of *Photorhabdus* are important to human health. The genomic differences between *P. asymbiotica* and other members of the genus appear to have arisen from recent horizontal transfer events, including the acquisition of *sopB* and other virulence factors (Tounsi *et al.*, 2006).

1.11 Objectives

The current study focuses on the biological and structural properties of a *Photorhabdus* protein first identified in the human pathogenic species *P. asymbiotica.* A two-dimensional gel electrophoresis of proteins in the culture supernatant of *P. asymbiotica* grown at 37°C, human body temperature, and at 28°C, closer to insect environmental temperature, identified a number of different proteins that may have specific functions in these niches (**Fig. 1.4**). Among those only expressed only at the lower temperature was a protein later named Mns, because of its apparent ability to make nematodes sticky.

In *P. luminescens*, high levels of Mns are secreted, and lower levels have been detected in cytoplasmic and membrane preparations (Turlin *et al.*, 2006). Secondary phase cells showed a dramatic down-regulation of Mns synthesis. The authors suggested the protein (referred to as plu1537) might be an insecticidal toxin, based on a weak similarity to the *B. thuringiensis* Cry34A toxin. Here, the *mns* gene has been cloned and expressed as a recombinant protein, and allowed to interact with *C. elegans* nematodes. Electron microscopy of the resulting nematode aggregates revealed a matrix-like material that coated the nematodes, particularly at the head and tail. We found it of interest to purify the protein to examine its structure and investigate its likely role in *Photorhabdus* biology.

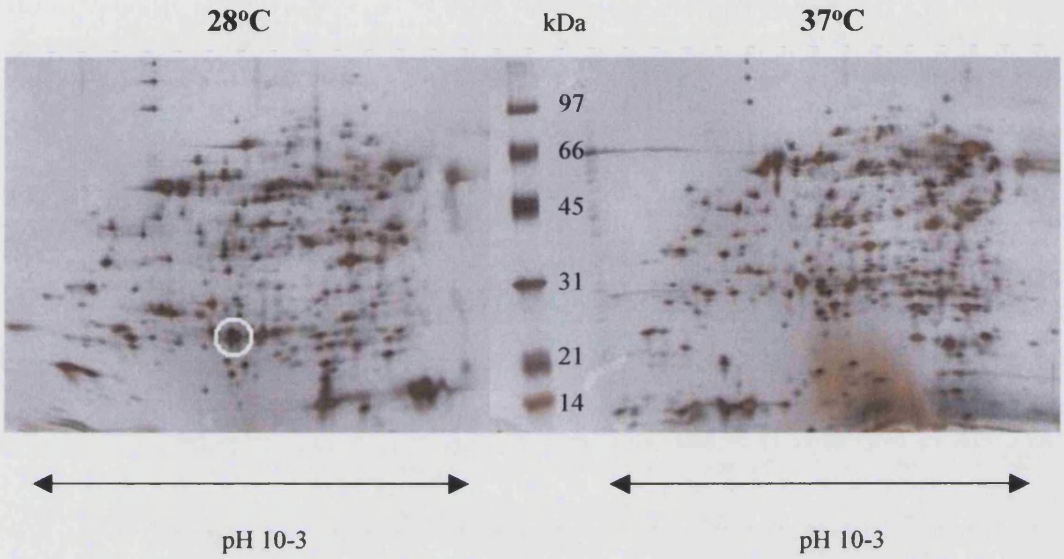


Fig. 1.4 *P. asymbiotica* supernatants separated by 2-D gel electrophoresis. *P. asymbiotica* cultures were grown at 28°C and 37°C, and supernatants were collected. The preparations were separated by isoelectric point along a pH gradient and, following denaturation, by mass. The two-dimensional gel identifies several proteins with differential expression in these conditions, including Mns (circled).

CHAPTER 2

The Biological Properties of Mns

2.1 INTRODUCTION

Mns is highly conserved among the *Photorhabdus* genus, and is highly expressed in *P. asymbiotica.* Two-dimensional electrophoresis data showed that Mns was not produced in 37°C, so is unlikely to be involved in infections, at least not in human infections. To examine the possibility that Mns is involved in symbiosis or other interactions with nematodes, *P. asymbiotica* culture supernatants were incubated with *C. elegans* and a nematode aggregation phenomenon was observed. The apparent ability of the supernatant to bind free nematodes together was further investigated, and lead to studies of cell attachment and biofilm formation on abiotic surfaces.

Mature biofilms were studied using transmission electron microscopy (TEM) of wild-type *P. luminescens* and an Mns-knockout strain *P. luminescens*, grown on agar to produce colonies. Biofilm development was also monitored, using cell attachment assays and confocal microscopy of biofilm structures forming on glass slides. The initial stages of attachment were examined with surface plasmon resonance (SPR).

2.1.1 Surface Plasmon Resonance

Surface plasmon resonance is a surface sensitive optical technique that can be applied to the study of biological interactions. The technique provides real-time data without the use of labels on both the affinity of the interaction and its kinetics, and has been used in the present study to monitor the early stages of *Photorhabdus* interaction with an abiotic surface during biofilm formation.

2.1.1.1 Principles of Surface Plasmon Resonance

In SPR, a beam of plane-polarised light is directed into a prism at an angle that results in total internal reflection; all light is reflected out of the prism and can be recorded by a detector. When the beam is directed onto a thin film of gold evaporated onto the reflective face of the prism, the light can interact with the free electrons of the gold and form plasmon resonance waves. These waves extend perpendicularly from the surface with a decay length of about 200nm (Lee *et al.*, 2005).

Light entering the prism at this angle of resonance is lost to the gold in the creation of the plasmon waves, and so less light is reflected. A detector records a dip in signal intensity due to this massive reduction in reflectivity. The angle of resonance is determined largely by the refractive index on the backside of the gold film, and if molecules are adsorbed onto this layer it causes a shift in the resonance to a higher angle (**Fig. 2.1**). It is therefore possible to monitor the process of absorption onto this surface in real time by measuring the shift in reflectivity.

The gold coating on the glass prism may be left bare or chemically modified to allow covalent binding of molecules, forming a uniform layer of immobilised ligands to which analytes can be introduced. A solution containing the analyte, such a DNA or proteins, is passed over the gold surface from a liquid handling system in a continuous flow (Szabo *et al.*, 1995). As analytes flow past the immobilised ligands, the molecules that bind and coat the surface shift the resonance angle, and the magnitude of the shift can be used to calculate the optical thickness of the coating. Furthermore, the association and dissociation rates of the interaction can be calculated (Knoll *et al.*, 2000, Malmqvist & Karlsson, 1997). However, in many cases a visual assessment of binding curves may be all that is necessary to characterise these events, particularly if one aims to compare the interactions of different analytes, such as different or mutated proteins. Lastly, samples containing the analyte need not be highly purified, and unknown proteins can be screened from crude samples (Torreri *et al.*, 2005).

2.1.1.2 Applications of Surface Plasmon Resonance

SPR is commonly used to study the absorption of thin films onto surfaces, DNA-protein and antibody-antigen interactions. Although real-time data can be collected, SPR has not been extensively used to monitor living systems, such as the kinetics of the first stages of bacterial adhesion to surfaces. Amano *et al.* (1999) have reported the use of SPR to examine the interaction between isolated fimbriae and surface-immobilised host proteins such as fibrinogen and haemoglobin, but investigations using whole cells as the analyte in an attachment study are lacking.

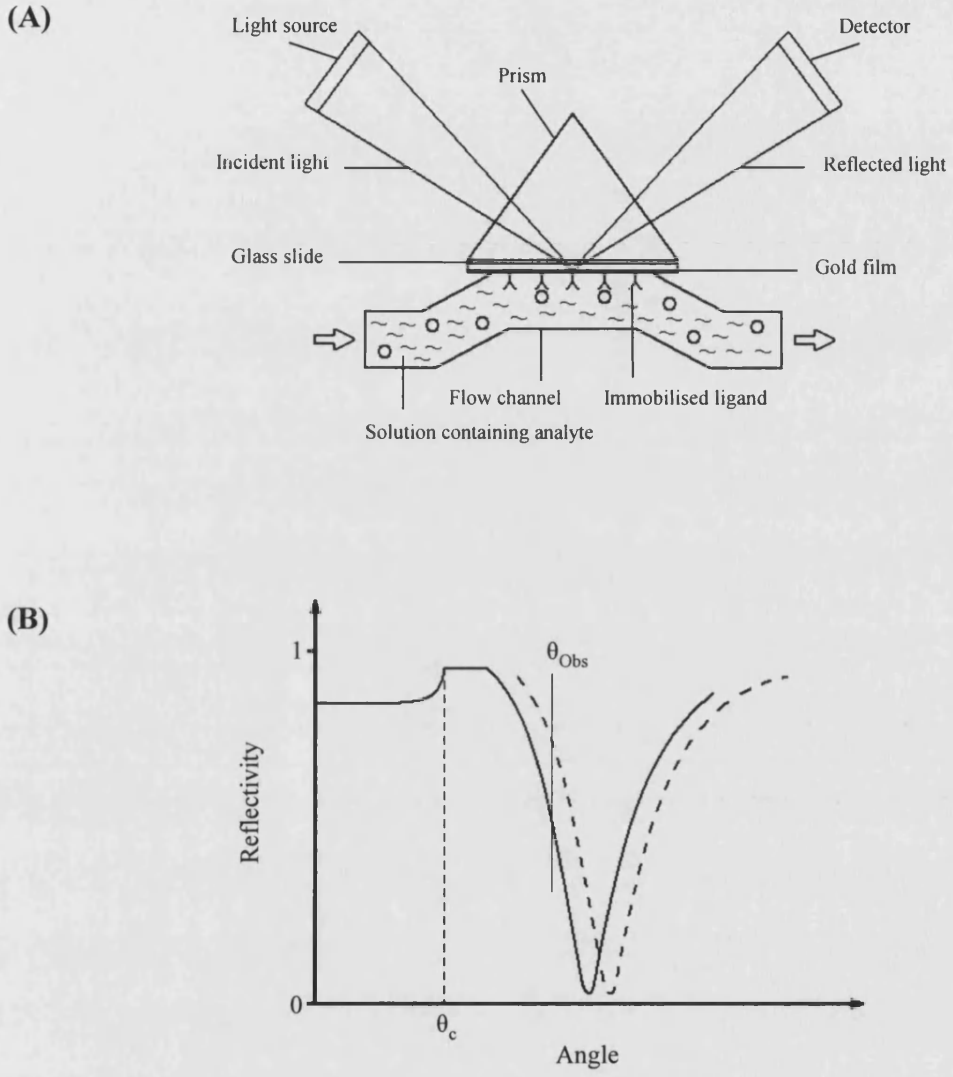


Fig. 2.1 Surface plasmon resonance configuration and reflectivity curve. **(A)** Schematic diagram of SPR instrument using Kretschmann configuration with liquid handling system. The incident light is directed through the prism to a thin gold film on an adjacent glass slide. Molecules passing over the glass slide attach to ligands immobilised on the gold surface and alter the angle of reflectivity that is recorded by the detector. **(B)** Typical reflectivity curve that dips due to surface plasmon resonance at the interface between the gold film and a solution. The resonance curve shifts to a higher angle (dashed line) when molecules attach to the surface. By measuring the reflected intensity at a fixed angle (θ_{Obs}), kinetic information on the attachment process can be monitored. The critical angle (θ_c), above which total internal reflectivity is achieved, is also shown. (Adapted from Knoll *et al.*, 2000).

Jenkins *et al.* (2004) have used *P. aeruginosa* as a model organism to explore this area. *P. aeruginosa,* an important human pathogen that often forms biofilms in persistent infections, is able to attach to gold surfaces, and SPR was able to detect attachment within seconds of the bacteria being introduced. Furthermore, the technique was shown to be sensitive enough to detect differences in attachment between wild-type and mutant strains, and between bare gold and a surface modified to increase electrostatic repulsion. Studies of this sort will contribute to the understanding of bacterial adhesion and colonisation, and may be used in the design of new antimicrobial surfaces and surface treatments.

SPR has been used in the current study to examine the early stages of attachment of *Photorhabdus* from growing cell cultures. Wild-type TT01 and an Mns knock-out strain grown to identical cell densities were tested, as well as supernatants of these cultures and cells washed in phosphate buffered saline (PBS) to identify differences between the strains and characterise these differences, with the aim of understanding the role of Mns in biofilm formation.

2.1.2 Characterisation of Mns

In addition to studies of biofilm formation, differences in the ability of wild-type and Mns-knockout *P. luminescens* to swim and to swarm were studied. Both forms of motility rely on functional flagella; swimming occurs in liquid media and in low-agar concentrations, while swarming occurs on surfaces and requires the production of surfactants to reduce surface tension. Swarming is widespread among flagellated bacteria and is a social phenomenon that involves the coordinated and rapid movement of cells. Furthermore, swarming cells differentiate from short and motile swimmer cells into longer and more highly flagellated cells. In *P. aeruginosa* a large number of genes influence swarming motility, and many mutants with impaired swarming motility also show defects in biofilm formation (Overhage *et al.*, 2007). Swarming and swimming motility was monitored in *Photorhabdus* on low agar concentration media, with wild-type and Mns-knockout strains of *P. luminescens*. To investigate the possibility that Mns is a secreted toxin of *Photorhabdus*, oral and injectable toxicity studies were performed using purified recombinant protein and wild-type verses Mns-knockout strain assays. Lastly, expression of Mns in both nematodes and insects was explored using Western blots analyses.

2.2 MATERIALS AND METHODS

2.2.1 Cloning and Expression of Mns

The *mns* gene was cloned into a pBAD30 vector (Guzman *et al.*, 1995) transformed into EC100 *E. coli* cells (Epicentre) and was the kind gift of G. Yang (University of Bath). Transformed cells were streaked from a frozen stock onto Luria Bertani lysogeny broth (LB) agar containing 100μg/ml ampicillin (amp) and grown overnight at 37°C. A single colony was picked and used to inoculate a 3ml overnight culture (LB amp), which was grown at 37°C with shaking at 180rpm. 1ml of this culture was pelleted (16,000x g, 2min) and used to inoculate 100ml LB amp, which was grown to an optical density at 600nm (OD_{600}) of 0.6-0.9, measured with a spectrophotometer (Spectronic Unicam) with a 1cm path length. The culture was induced to express Mns by addition of a sterile 10% L-arabinose (Sigma) solution to a final concentration of 0.2%. A 1ml sample was taken before induction, at hourly intervals (to 4h) and after overnight growth to monitor expression.

These samples were pelleted and the cells were resuspended in 100μl of 1x SDS (sodium dodecyl sulphate) loading buffer (Sambrook *et al.*, 1989). Proteins in the supernatant were precipitated with trichloroacetic acid (TCA); 100μl ice-cold 100% TCA (w/v) was added to 1ml supernatant and vortexed for 15 seconds. After overnight incubation at 4°C, the suspension was centrifuged at 16,000x g, 10min at 4°C. The supernatant was discarded and the pellet resuspended in 100μl PBS followed by 500μl acetone, and centrifuged at 16,000x g, 5min, room temperature (RT). The resuspension and acetone wash was repeated; the pellet dried in air and resuspended in 100μl 1x SDS buffer. 10μl was loaded to a 10% SDS-polyacrylamide gel for electrophoresis (SDS-PAGE) (Sambrook *et al.*, 1989), along with the resuspended pellets (amount to load calculated as reciprocal of OD_{600} x 4, in μl), to determine when Mns was optimally expressed and if the product was secreted. SDS gels were run at 12Amps constant current/80-100V, and were stained in Coomassie blue dye overnight. Gels were destained in 10% acetic acid, 30% methanol, 60% water and fixed with 10% acetic acid, 30% methanol, 55% water, 5% glycerol, before drying in cellulose membranes (Biorad).

The pellet of 100ml of *E. coli* strain EC100 expressing Mns overnight at 37°C was resuspended in 10ml of buffer A (20mM HEPES pH 7.5, 50mM NaCl) and the cells were lysed with 8 x 10s pulsed sonication in a Branson digital sonifer sonicator. The lysed cells were pelleted at 16,000x g, 20min, 4°C. The pellet was resuspended in 10ml buffer A, and 10µl was loaded onto an SDS-polyacrylamide gel as described above. 10µl of the supernatant was also loaded on the gel, alongside a whole-cell fraction to determine the solubility of the protein.

2.2.2 Purification of Recombinant Mns

The pellet of 1l of cells expressing Mns was lysed as described above and the supernatant was diluted to 20ml with buffer A. The lysate was loaded as 5ml fractions onto a 5ml Hitrap QFF anion exchange chromatography column (GE Healthcare) equilibrated with 3x column volumes buffer A, 3x column volume buffer B (20mM HEPES pH 7.5, 1M NaCl), 3x column volumes buffer A. All chromatography was performed on an ÄKTApurifier (GE Healthcare). The column was run at 0.8ml/min with a 15ml column wash after loading and a 5 column volume gradient from 5% to 100% buffer B to elute the protein. This was later optimised to a stepwise elution with a step to 20% buffer B, 5 column volume gradient to 70% buffer B, step to 100% buffer B. The column was washed with 5ml buffer A before reloading. 1ml fractions were collected in 1.5ml microcentrifuge tubes, and 10µl samples were loaded for SDS-polyacrylamide gel electrophoresis.

The Hitrap QFF chromatography was repeated with fractions containing Mns to improve the purity. The fractions were diluted fourfold with buffer A to reduce the NaCl concentration before loading in 5ml fractions.

For crystallisation, the first Hitrap QFF step was followed by anion exchange with a 1ml MonoQ column (GE Healthcare) and a Superdex 75-10 (GE Healthcare) gel permeation column. For MonoQ, fractions containing Mns were diluted and 4ml was loaded onto the column after equilibration as above. The column was washed with 8ml buffer A at 0.8ml/min, a further 4ml was loaded and the column wash was repeated. Mns was eluted with a gradient of 5%-25% buffer B over 8 column volumes, and fractions containing Mns were identified by SDS-PAGE.

For gel permeation, fractions containing Mns were concentrated in a 5,000 MWCO centrifuge filter concentrator at (Millipore) at 800x g to a volume of 300µl. The protein sample was centrifuged at 4,500x g, 4min in a microcentrifuge tube prior to loading 30µl onto the column, which was equilibrated at 0.25ml/min with 20mM HEPES pH 7.0, 150mM NaCl. The column was run at 0.25ml/min and 0.5ml fractions were collected. The purity of Mns was analysed by SDS-PAGE.

Protein concentration was determined spectrophotometrically at 280nm. Purified protein was denatured by 100-fold dilution in 6M guanidine hydrochloride and readings were taken in a 1cm path length quartz cuvette. Using Beer's law, the molar concentration was calculated as OD280/(molar extinction coefficient x pathlength (cm)). The molar extinction coefficient was calculated using the ProtParam program (ExPASy server, www.expasy.org, Gasteiger *et al.*, 2005).

2.2.3 Purification of Native Mns

To purify native Mns from *P. asymbiotica* supernatant, a 200ml LB culture was grown at 28°C for 48h from a single colony. Cells were pelleted by centrifugation at 15,000x g, 20min, 4°C. The supernatant was diluted 1:4 with 20mM HEPES pH 7.5 and filtered through a 0.22µm filter (Millipore) using a vacuum pump. The filtered supernatant was loaded onto a 5ml Hitrap QFF column (GE Healthcare) and an elution gradient using 20mM HEPES pH 7.5, 1M NaCl was used to separate the proteins.

2.2.4 Nematode Interaction Study

P. asymbiotica (strain ATTC43949) was streaked from a frozen stock on LB agar at 28°C overnight. A single colony was picked and used to inoculate a 5ml LB broth culture, shaking at 180rpm at 28°C. A 1ml sample of this culture was pelleted at 16,000x g, 2min in a microcentifuge tube and the cell pellet was used to inoculate flasks containing 100ml LB. These cultures were grown at 28°C and 37°C for 2 days before pelleting at 15,000x g, 10min, 4°C.

Culture supernatants were applied neat to wells of a 25-well plate (Sterilin) and in dilutions with LB. *C. elegans* var. Bristol (strain N2) grown on *E. coli* were harvested from an agar plate into 4ml of LB, and 100µl was carefully pipetted into each well. Activity in the wells was monitored with a Nikon SMZ-1500 stereomicroscope. Samples of culture supernatant were also incubated at 4°C, RT, 28°C and 37°C without agitation for 6-12 days before applying to the nematodes.

Nematodes aggregated with *P. asymbiotica* supernatant were isolated and resuspended in 30µl 1x SDS buffer. The sample was boiled at 100°C and 10µl was loaded onto a 10% SDS-polyacrylamide gel. The presence or absence of Mns in the aggregation was detected by Western blot analysis. Western blots were performed following the directions of Sambrook *et al.*, (1989). Briefly, an SDS-polyacrylamide gel was placed in Western transfer buffer (14.4g glycine, 3.03g Tris base, 200ml methanol, made to 1l with water) for 5min prior to placing onto pre-soaked Whatman paper. A nitrocellulose filter (Biorad) also soaked in Western transfer buffer was lowered on top, followed by Whatman paper. The assembly was sandwiched between foam sheets in a Biorad gel tank, which was filled with ice-cold Western transfer buffer. The blot was performed over 60min at 100V. The nitrocellulose filter was immediately placed in 50ml phosphate buffered saline (PBS), milk powder (5% w/v), tween (0.1% w/v) (TPBS) and kept at 4°C overnight.

The nitrocellulose filter was placed in 25ml TPBS and incubated with gentle agitation at RT for 90min with 5µl Mns monoclonal rabbit antibody (provided by N.R. Waterfield). The antibody was washed off with 3x 10min 50ml TPBS before placing in fresh TPBS with 40µl anti-rabbit antibody conjugated with alkaline phosphatase (Sigma). After 90min at RT with gentle agitation, the secondary antibody was washed off with 2x 50ml TPBS and 2x 50ml PBS. Excess buffer was drained off and antibody was visualised with 175µl of NBT-BCIP (Fluka) in 10ml NBT-BCIP buffer (100mM Tris, 10mM NaCl, 5mM $MgCl_2$, adjusted with dilute HCl to pH 9.5). When bands became visible, this buffer was washed off with water.

E. coli expressing Mns were also used in nematode interaction studies, and were prepared as described above. Cells were grown overnight at 37°C after induction with L-arabinose. To detect Mns in association with nematodes, *C. elegans* were collected from a plate with 2ml of PBS and the nematodes were allowed to settle in microcentrifuge tubes. The PBS was aspirated off to leave 50µl suspended nematodes. 20µl of purified Mns at 10mg/ml was made up to 50µl with PBS and added to the nematodes and left to incubate for 15min. Negative controls was prepared without nematodes, and with nematodes incubated with PBS only.

The nematodes were washed with 1ml PBS after pelleting at 16,000x g. The supernatant was removed, the worms resuspended in 1ml PBS and the tube was re-centrifuged. This step was repeated and the supernatant was aspirated off. 30µl of 1x SDS buffer was added to the nematodes. The no-nematode and no-Mns controls were prepared in the same way and 10µl from each tube was loaded for SDS-PAGE after heating at 100°C. Kingscliff nematodes, the host strain of Kingscliff strain *P. asymbiotica*, were also suspended in 30µl 1x SDS buffer for SDS-PAGE. Western blot analysis was performed as above.

2.2.5 Microscopy of Nematode Aggregates

Nematodes aggregated by Mns-expressing *E. coli* were prepared for scanning electron microscopy (SEM) with a JEOL JSM6310 microscope at the Centre for Electron Optical Studies (University of Bath) with the help of U. Potter. The *E. coli* were also examined by transmission electron microscopy (TEM).

2.2.6 Wild-Type and Mns-Knockout *P. luminescens* TT01

An Mns-knockout strain of *P. luminescens* TT01 was kindly provided by G. Yang (University of Bath). In brief, the knockout mutant was created by polymerase chain reaction (PCR) amplifying the regions upstream and downstream of the *mns* gene. The fragment was ligated into the suicide vector pDS132 for allelic exchange with wild-type *P. luminescens* TT01. The gene deletion was confirmed by PCR and precipitation of culture supernatants that showed an absence of Mns compared to wild-type supernatant.

Both the wild-type strain of *P. luminescens* TT01 and the Mns-knockout strain were rifampicin (rif) resistant (50μg/ml rifampicin) after selection on increasing concentrations of antibiotic. Mns-knockout TT01 was used in cell attachment assays, both to polypropylene for classical biofilm analyses and to gold and modified gold surfaces for SPR. Mns-knockout TT01 was also used in comparisons of biofilm structure by transmission electron microscopy, and in pellicle formation.

P. luminescens TT01 (wild-type and Mns-knockout) and *P. asymbiotica* strains were also made to express green fluorescent protein (GFP) constitutively from the tetracycline resistant pHC60 plasmid (TT01) and the kanamycin resistant pBAMH7 plasmid (*P. asymbiotica*). These cells were used for imaging *Photorhabdus* biofilms growing on glass slides by confocal microscopy.

2.2.7 Mns Expression in *Photorhabdus* Cultures

To follow expression of Mns in *Photorhabdus*, 1ml of a *P. asymbiotica* overnight culture was used to inoculate 100ml LB. This was shaken and divided into 2x 50ml cultures in 250ml flasks. One culture was grown at 28°C and the second at 37°C, with 250rpm rotation.

From the growing cultures at regular intervals 1ml samples were taken and the optical density was measured. The cells were pelleted at 16,000x g, 5min in a microcentrifuge. The supernatants were chilled on ice before being treated with TCA (as above). 5-15μl of the TCA preparation, resuspended in 1x SDS buffer, was loaded for SDS-PAGE and a Western blot was performed.

2.2.8 Mns Expression During Nematode Association and Insect Infection

P. asymbiotica strain ATTC43949 and Kingscliff *P. asymbiotica* were grown overnight in LB and the cells of a 1ml sample were pelleted by centrifugation at 16,000x g, 5min. The supernatant was treated with TCA (as above) to precipitate the secreted proteins. Expression of Mns from these strains was analysed by Western blot analysis, along with crushed Kingscliff nematodes, the host vector of Kingscliff *P. asymbiotica.* Nematodes containing the bacteria were crushed in PBS after passage through *G. mellonella.* 10µl of crushed nematode and 10µl of TCA preparations were mixed with 3µl 3xSDS gel loading buffer and loaded for SDS-PAGE. Recombinant Mns was also loaded as a positive control. The gel was transferred to a nitrocellulose filter and monoclonal anti-Mns antibody was used to detect the protein.

To examine Mns expression during infection, 10 *G. mellonella* were infected by injecting each with *P. luminescens*, *P. asymbiotica* strain ATCC43949 or Kingscliff *P. asymbiotica.* The cultures were prepared as overnight cultures and diluted to an optical density (600nm) of 0.1, and 20µl was injected into each insect. Insects were collected on successive days and crushed in 4M urea, 2% SDS, 5% β-mercaptoethanol, 125mM Tris pH 8 (1ml/insect). From this 10µl was taken and samples were loaded for SDS-PAGE and Western blot analysis as above.

Day zero fifth instar *M. sexta* larvae were injected with 50µl wild-type *P. luminescens* TT01, Mns-knockout TT01, or *P. asymbiotica* ATTC43949 overnight cultures diluted to OD_{600} 0.1. After 24h, the haemocytes, fat body and midgut were isolated from two insects and crushed using a plastic pestle in 500µl of *Manduca* saline buffer (4mM NaCl, 40mM KCl, 18mM $MgCl_2.6H_2O$, 3mM CaCl, 1.7mM PIPES, 50g/l sucrose, 1g/l polyvinyl pyrollidone, adjusted to pH 6.5 with 5N KOH, (Willott *et al.*, 1994)). The tissue was pelleted and 80µl of the supernatant was mixed with 20µl 5xSDS gel loading buffer with 5% β-mercaptoethanol. 10µl was loaded for SDS-PAGE and Western blot analysis.

2.2.9 Mns Toxicity

E. coli containing *mns*-pBAD30 and empty vector pBAD30 were grown in 50ml LB cultures and induced overnight. The cells were pelleted and resuspended in 10ml PBS prior to sonication. 10µl was injected into *G. mellonella*, which were incubated at RT.

Heterologously expressed Mns was purified with two steps of anion exchange chromatography as described above. The protein was concentrated to 0.5mg/ml, and serially diluted in 20mM HEPES pH 7.5, 0.6M NaCl, the buffer conditions that eluted the protein, before injecting into *G. mellonella*. This buffer was also used as a negative control. As a protein control, bovine serum albumin (BSA) (New England Biolabs) concentrated at 0.5mg/ml was injected. 10µl was injected into each insect, which was incubated at 25°C overnight.

To assess any oral toxicity, $1cm^3$ artificial wheat germ food blocks were supplemented with 100µl 0.5mg/ml purified Mns or BSA, or with 100µl of supernatant of lysed *E. coli* cultures induced or not induced to express from the *mns*-pBAD30 vector. The bacterial cultures were applied neat, diluted 1:10 or diluted 1:100 with phosphate buffered saline (PBS), and all supplements were allowed to dry for approximately 20min under laminar flow. Two first-instar *M. setxa* neonate larvae were placed on each food block and after 7 days incubation at 25°C the weight of each caterpillar was recorded.

2.2.10 Biofilm Formation on Polypropylene Plates

Wild-type *P. luminescens* TT01 and Mns-knockout TT01, both rif resistant, were diluted from overnight cultures to an OD_{600} of approximately 0.06, using LB rif (50µg/ml). 125µl of the diluted suspension was pipetted into individual wells of 96 well polypropylene plates (Corning Incorporated). *P. asymbiotica* strain ATCC43949 and *E. coli* containing the *mns*-pBAD30 vector were similarly prepared using LB amp supplemented with 0.2% L-arabinose for expression from *E. coli*, which were pipetted into polyvinyl chloride plates (Becton Dickinson).

The 96 well plates were covered with a loose-fitting lid and placed on a tray with wet paper towels to prevent dehydration. The plates were grown for 48h without agitation in a sealed container, at 28°C for *P. luminescens* and 28°C and 37°C for *P. asymbiotica* and *E. coli*. After incubation, biofilm formation was assessed according to the method used by Ferrières and Clarke (2003). Briefly, the loosely bound cells were gently aspirated off and wells were washed with 200µl water. 200µl 1% (w/v) crystal violet (CV) was added to each well and incubated for 5min. The CV was removed and the wells were washed with water until the water showed no colour. 200µl acetone-ethanol (20:80) was added to dissolve the CV that had stained the attached cells. 100µl of this solution was removed and added to 900µl water. The optical density was recorded at 570nm as a measure of the extent of biofilm formation in each condition.

2.2.11 Pellicle Formation Assay

To monitor biofilm formation over an extended time period and allow observation of pellicles, 1µl of overnight *P. luminescens* TT01 wild-type (rif resistant) and Mns-knockout cultures were inoculated into 10ml LB and grown statically at 28°C. Pellicles were photographed every 24h for 9 days. The assay was performed in triplicate.

2.2.12 Transmission Electron Microscopy of *Photorhabdus* Colonies

P. luminescens TT01 and Mns-knockout TT01, and *P. asymbiotica* strain ATCC43949 were grown from a frozen stock for 3 days on LB agar. Colonies were viewed with a JEOL JEM1200 transmission electron microscope (TEM) at the Centre for Electron Optical Studies (University of Bath) with the help of U. Potter. Slides were also prepared with Mns antibody and an anti-rabbit gold-conjugated secondary antibody to locate Mns within the biofilm structures.

2.2.13 GFP-*Photorhabdus* Attachment to Glass Slides

Photorhabdus strains expressing green-fluorescent protein (GFP) were streaked from frozen stocks on selective agar (50μg/ml rifampicin with 10μg/ml tetracylcin (tet) for *P. luminescens*, and 50μg/ml rif and 5μg/ml kanamycin (kan) for *P. asymbiotica*). Single green colonies were picked and grown with selection in 10ml LB broth for 2 days at 28°C without agitation (28°C and 37°C for *P. asymbiotica*). These cultures were diluted with LB to an OD_{600} of 0.1 and 500μl was pipetted into the wells of 12-well tissue culture plates (Corning Incorporated). Each well contained a sterile circular glass coverslip, and 12 wells were filled to allow two glass slides to be taken at designated time points. The plates were sealed with parafilm and grown without agitation. *P. luminescens* cultures were grown at 28°C and *P. asymbiotica* was grown at 28°C and 37°C. After 12h the first cultures were aspirated and the glass coverslip was washed carefully by application of 3x 1ml of 0.9% NaCl solution (Ploux *et al.*, 2007). The glass slide was fixed by application of 300μl of 4% paraformaldehyde to each well. The plate was left at RT for 20min and at 4°C for 24h. Finally, the coverslips were removed from the wells, excess paraformaldehyde was drained off and the coverslips were inverted onto a drop of vectashield (Vector Laboratories) on a glass slide, such that the attached cells were sandwiched between the glass. GFP-expressing cells were viewed with a confocal microscope (Nikon Eclipse 90i).

The expression of Mns during the experimental time period was monitored by performing a Western blot on the aspirated cell cultures. The cell pellet and TCA-precipitated supernatant were prepared for SDS-PAGE and Western transfer as described above.

2.2.14 Surface Plasmon Resonance

2.2.14.1 Bacterial Strains and Preparation

Wild-type *P. luminescens* TT01 were grown to an optical density (600nm) of 1.6 and 2x 1ml was pelleted at 16,000x g, 5min and the supernatant was aspirated off. To study proteins bound to the cell membrane, the cell pellet was resuspended in PBS, pelleted and finally resuspended in 1ml PBS. This created a suspension of cells in a protein-free supernatant. The cells were serially diluted in PBS from 0.59E9 colony forming units (CFU)/ml to 0.59 CFU/ml, to establish appropriate cell densities for repeatable SPR measurements.

2.2.14.2 Instrumentation and Measurements

SPR measurements were measured on a home-built SPR instrument based on the design of Knoll. Light at a wavelength of 633nm from a HeNe laser was plane polarised and directed through a LaSFN9 (Berliner glas) prism onto the underside of a gold disk (Fluoroware) using the Kretschmann configuration. Angle reflectivity plots were obtained before and after adsorption of bacterial cells/bacterial proteins. The prism was rotated to a fixed point on the SPR curve, around 1° negative of the resonant minimum, and the increase in reflectivity was measured in real time as material adsorbed to the surface.

2.2.14.3 Experimentation

Cell and supernatant preparations were added to the SPR gold disk in 50µl applications. Initially, 50µl of PBS-washed *P. luminescens* cells were added from channel 1, with 50µl PBS applied from channel 2, and measurements were taken over 1000 s from each of the dilutions in the series of 0.59E9 CFU/ml to 0.59 CFU/ml. The gold substrate disk was lightly rinsed with distilled water, allowed to dry and stained with 1% crystal violet dye for 5 min. The substrate was rinsed thoroughly with distilled water before observation under a light microscope (Jenkins *et al.*, 2005).

To make comparisons between the attachment of TT01 and Mns-knockout, cells of both strains were washed in PBS and concentrated as described above. Supernatants from these strains were also collected and analysed by SPR. The experimental design would therefore allow comparisons to be made between cells potentially expressing Mns on their outer membrane, and those unable to express the protein. Comparisons could also be made between the attachment of extracellular material containing secreted Mns and not containing Mns (supernatant from Mns-knockout). Whole-culture preparations were also analysed to compare attachment of cells in their native supernatant, and the attachment of washed *E. coli* strain E100 induced to express recombinant Mns was compared with an EC100 control.

In addition to an unmodified gold surface, TT01 attachment was examined on a gold surface with a self-assembled monolayer (SAM) designed to alter the hydrophobicity of the surface. Octadecylmercaptan (ODM) was immobilised on a gold surface by evaporation; a gold SPR disk was placed in a beaker containing 5ml of 1mM ODM in chloroform and allowed to evaporate overnight. PBS-washed *P. luminescens* TT01 cells diluted to 4.9E6-4.9E8 cfu/ml were applied to the modified disk for comparison with the more hydrophilic unmodified gold. A gold disk was also treated with an intermediate surface, mercaptopropionic acid, again by evaporation. Furthermore, the ODM SAM was removed by 15min treatment with ozone and SPR readings were repeated.

2.2.15 Swimming and Swarming Assays

Swimming plate medium was prepared with 0.1% LB broth and 0.25% agar. Plates were inoculated by stabbing into the centre of the semi-solid agar with growth from LB agar plates. Swarming assays were performed using 0.1% LB broth and 0.45% agar. The plates were kept at 28°C and monitored after 24h.

2.3 RESULTS

2.3.1 Nematode Interaction Study

Incubation of *C. elegans* with *P. asymbiotica* cultures has been associated with an aggregation phenotype of the nematodes (N.R. Waterfield, unpublished observation). To investigate this phenomenon further, cultures and supernatants were serially diluted in LB and the aggregation phenotype was monitored by low magnification light microscopy. We found that clumping could be achieved with supernatant alone, and was not produced from *P. asymbiotica* grown at 37°C.

The aggregation phenotype was particularly pronounced around nematode eggs, with several nematodes assembling around a single egg. Furthermore, the nematodes were frequently held together at their tails, with aggregated material visible in the nematode mass from which they were unable to escape (**Fig. 2.2**). The phenotype was evident within ~15min of application of worms to the supernatant. Many worms became paralysed or showed much less motility within 1h. We also found that incubation of the supernatant at 4°C enhanced the phenotype. Paralysis and subsequent death of the nematodes was associated with cultures grown for extended periods (approximately three days or more). This effect appears to be separate from the clumping phenotype, which was detectable in low levels from early cultures.

Two-dimensional gel electrophoresis is currently being used to identify proteins secreted into the supernatant from cells grown at 37°C and not at 28°C, and therefore may be important in human infections by *P. asymbiotica*. Because the nematode aggregation phenotype was not associated with cultures grown at the 37°C, the two-dimensional gel was analysed for proteins expressed only at the lower temperature. The Mns protein was identified by mass spectrometry and the gene was cloned for expression by *E. coli*.

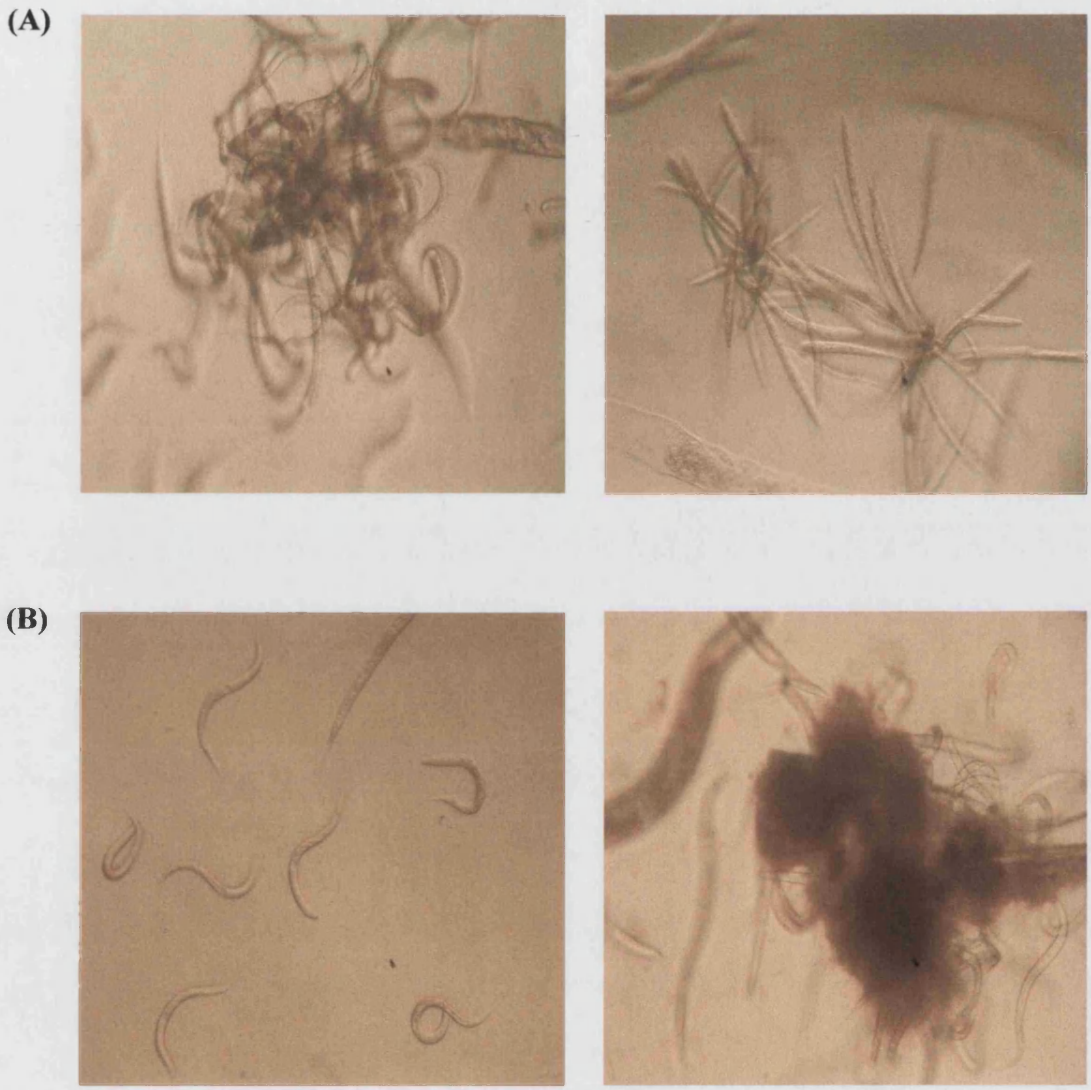


Fig. 2.2 Nematode aggregation phenotype. **(A)** Aggregation by *P. asymbiotica* supernatant does not kill *C. elegans* (left), but nematodes are paralysed by a second factor. This factor is produced by *P. asymbiotica* culture after prolonged incubation and results death of the nematodes (right). **(B)** Mns expression in *E. coli* EC100 reproduces the aggregation of *C. elegans* observed with *P. asymbiotica* supernatant. (left, uninduced cells; right, induced expression).

2.3.2 Cloning and Heterologous Expression of Mns

The *mns* gene was cloned into the pBAD30 vector and was expressed from EC100 cells by supplementation with L-arabinose to 0.2% final concentration. Mns was expressed within 4h of induction with L-arabinose, but was routinely expressed overnight to obtain a high yield of recombinant protein. Similar expression was observed from BL21 cells.

2.3.3 Purification of Recombinant Mns

The predicted amino acid sequence of Mns was analysed using the ProtParam program, which showed the calculated isoelectric point (pI) of the protein to be 4.83. This point was plotted among the pI of *E. coli* proteins to show that Mns could be purified using anion exchange chromatography, and that the majority of *E. coli* proteins, from which the protein would be purified, had a higher net positive charge than Mns (**Fig. 2.3**). A sample containing Mns was passed through a Q sepharose column. The target protein was captured while most of the contaminants washed through. To elute Mns, the salt concentration of the buffer was increased (**Fig. 2.4**). After dilution to reduce the salt concentration of the eluate, fractions were reloaded onto the column and the purity achieved was sufficient for NMR preparations. For crystallisation, a MonoQ column with finer sepharose resin was used, and the protein was polished with a gel permeation column (**Fig. 2.5**).

2.3.4 Purification of Native Mns

P. asymbiotica supernatant was diluted 1:4 with 20mM Hepes pH 7.5 and loaded onto a Q sepharose column for purification. Mns was collected in fractions and was visible on a SDS-polyacrylamide gel but was very dilute (not shown). We were unable to concentrate the protein further in a centrifuge concentration device due to apparent adhesion of the protein to the device filter. Fractions collected from the sepharose column were also applied to *C. elegans* to observe which fractions induced nematode aggregation. Fractions A9-A13 all induced clumping. Fractions collected after this caused no clumping but the nematodes died, likely as a result of the increasing salt concentration (**Fig. 2.6**). Interestingly, no nematode paralysis was observed, while the whole supernatant preparation caused both nematode aggregation and paralysis. This indicates the two phenomena are caused by different factors.

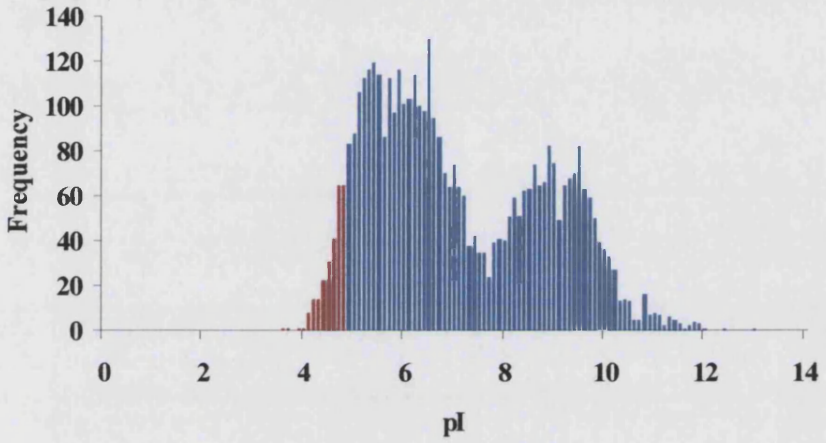


Fig. 2.3 Isoelectric point distribution of *E. coli* proteins. Histogram of pI values at 0.1 unit intervals for *E.coli* proteins shows a biomodal distribution, with the majority of proteins having a net positive charge higher than that of Mns (predicted pI 4.83) in a neutral buffer. This feature was exploited for the purification of recombinant Mns away from *E. coli* proteins. Red bars show the frequency of proteins with a pI less than Mns and blue bars represent proteins with pI values above 4.83. (Adapted from Schwartz *et al.*, 2001).

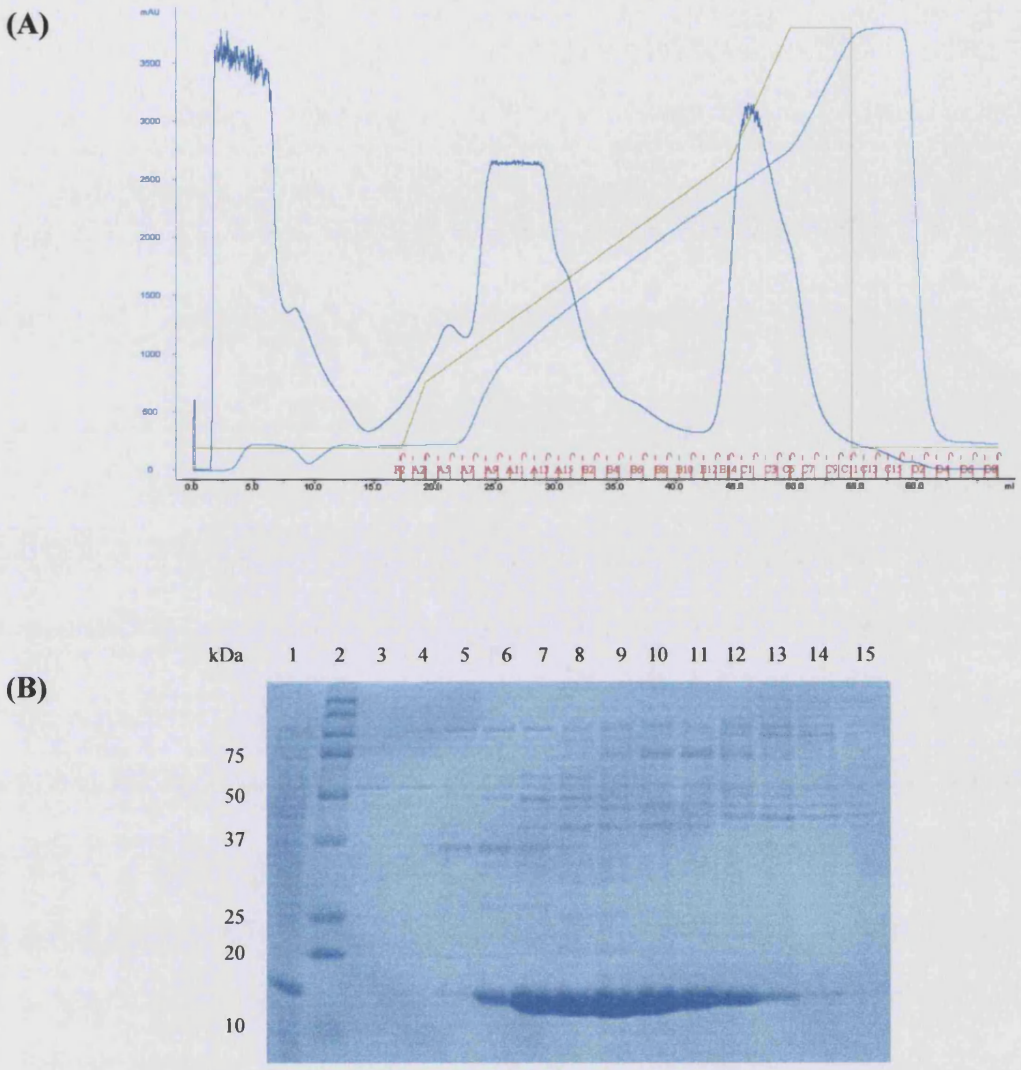


Fig. 2.4 First-stage purification of Mns. **(A)** Chromatogram shows Mns elution from HiTrap Q FF ion exchange column as salt concentration was increased from 5 to 70%. The 1ml eluted fractions were analysed by SDS-PAGE and those containing Mns were pooled for subsequent purification procedures. Navy line represents protein (218nm); green line is percentage Hepes buffer with 1M NaCl, light blue line is conductivity and red divisions are fractions. **(B)** 10% SDS PA gel shows Mns in fractions collected from ion exchange column. Lane 1, whole soluble fraction; 2, marker; 2-15, 1ml fractions eluting after increasing the salt concentration. Fractions shown in lanes 7-12 were pooled for further purification.

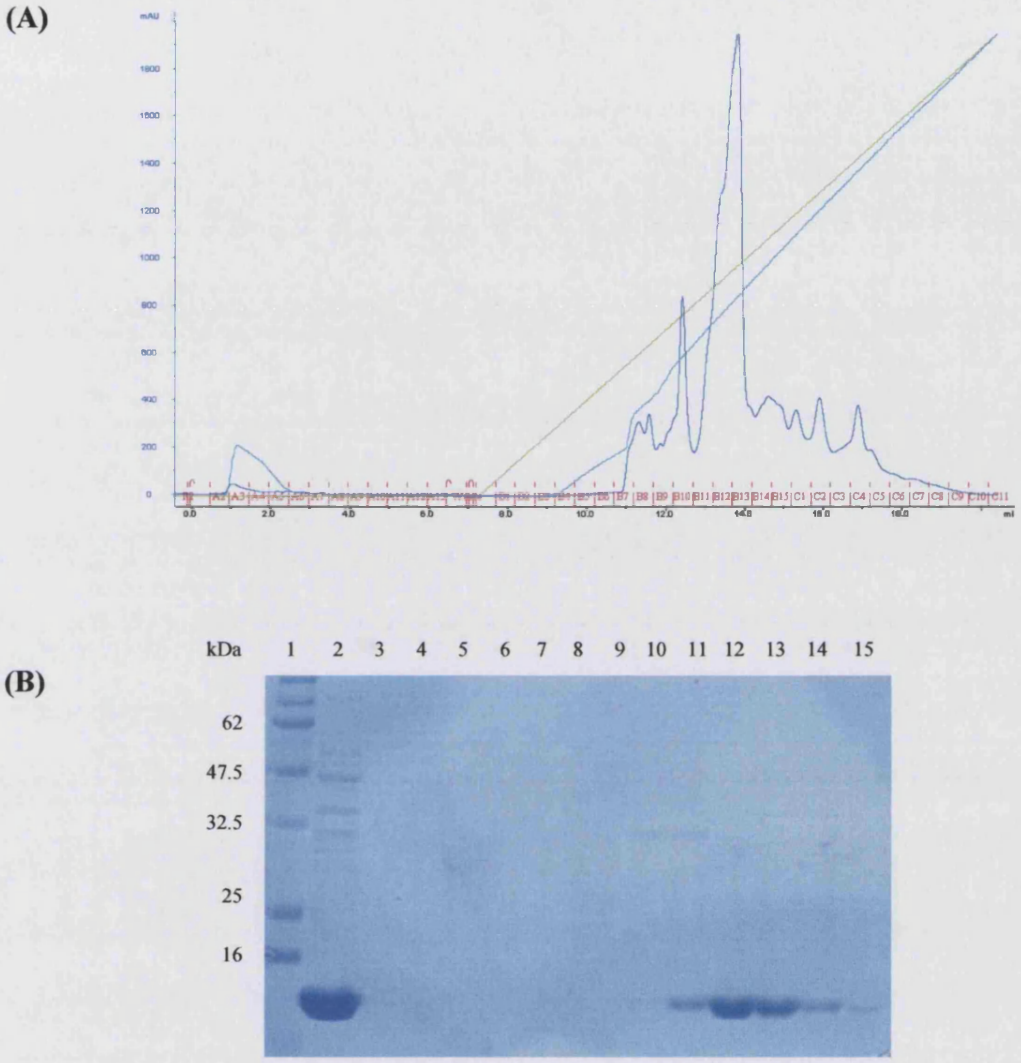


Fig. 2.5 Second-stage purification of Mns. Partially purified Mns was passed through a 1ml MonoQ column. **(A)** The chromatogram shows the separation of MNS (major peak) from *E. coli* proteins that contaminated the fractions obtained from the first-stage purification. Fractions containing Mns were pooled and diluted before loading onto the column, and the eluted protein was at sufficient purity for NMR. Navy line represents protein (218nm); green line is percentage Hepes buffer with 1M NaCl, light blue line is conductivity and red divisions are fractions. **(B)** SDS-polyacrylamide gel of eluted fractions, showing a highly pure Mns product. Lane 1, maker; 2, pooled fractions from first-stage purification; 3-15 fractions of MonoQ Elution, with Mns in lanes 11-15.

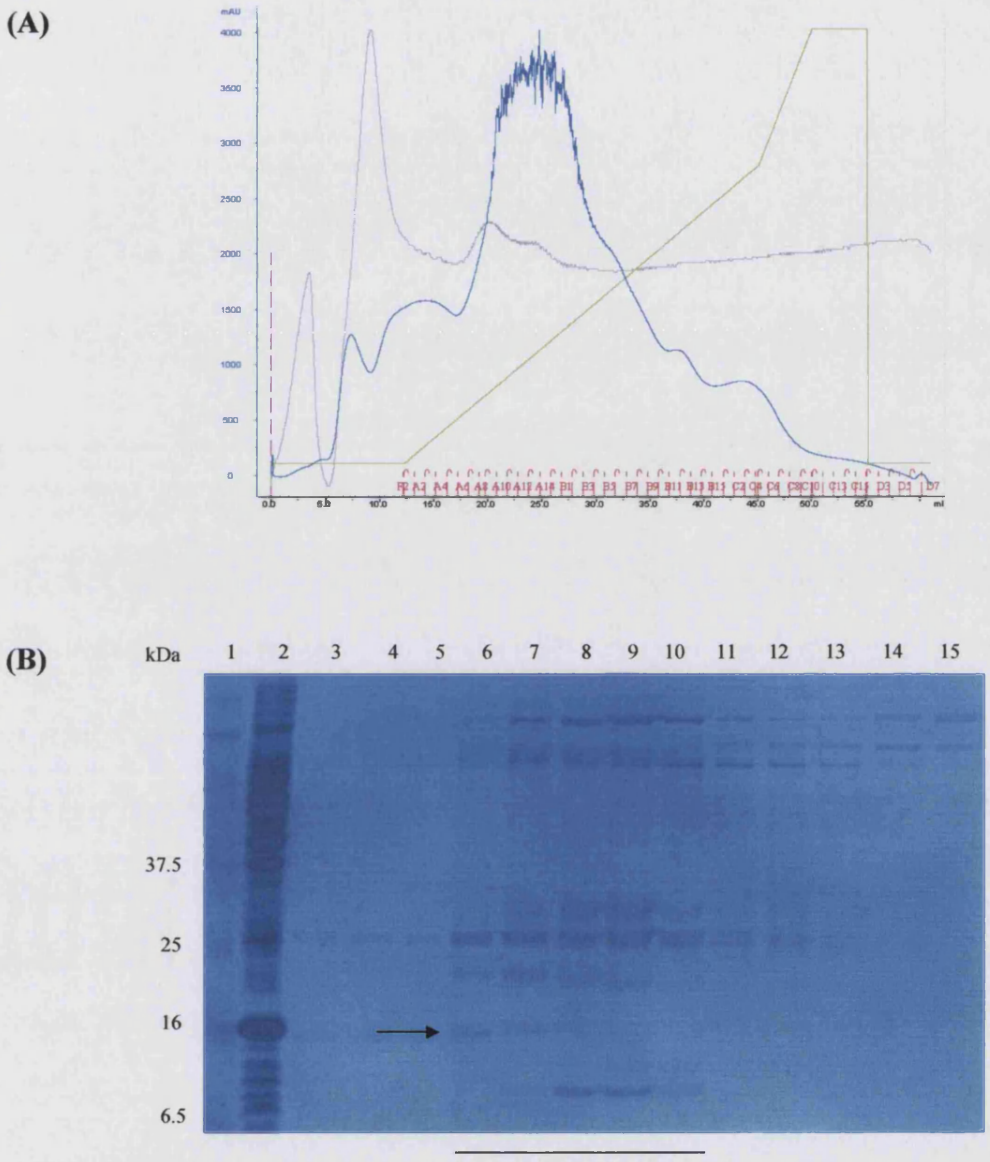


Fig. 2.6 Purification of native Mns. **(A)** Elution profile of *P. asymbiotica* supernatant purification. Blue line shows protein elution, measured at 280nm. *P. asymbiotica* culture supernatant was filtered and passed through a Q sepharose column. **(B)** Fractions were loaded onto an SDS-polyacrylamide gel. Mns could not be concentrated further using a centrifuge filter concentrator, likely due to adhesion to the filter. Fractions A9-A13 from the elution profile (underlined in lanes 6-10) were shown to aggregate nematodes and appear to contain a low level of Mns (arrow). Lane 1, protein marker; 2, recombinant Mns; 3, fraction A2; 4, A4; 5, A6; 6, A9; 7, A10; 8, A11; 9, A12; 10, A13; 11, A15, 12, B1; 13, B2; 14, B4; 15, B6.

2.3.5 Mns Association with Nematodes

A Western blot was performed using anti-Mns peptide rabbit polyclonal antibody to identify the presence of Mns in *C. elegans* aggregated by *P. asymbiotica* supernatant. The clumped worms were washed in PBS and boiled in loading buffer, and a band was clearly visible on the nitrocellulose blot, providing evidence that Mns is associated with the aggregated mass (**Fig. 2.7 (A)**). *In situ* histochemical detection was also used to verify the presence of Mns in these aggregations induced by *P. asymbiotica* supernatant. A fluorescent anti-rabbit secondary antibody highlighted the location of the anti-Mns peptide rabbit antibody; the aggregation was viewed with a confocal microscope and the green fluorescence was highly visibly, strongly indicating the presence of Mns (**Fig. 2.8**).

Expression of Mns in *E. coli* appeared to reproduce the aggregation phenotype observed with *P. asymbiotica*, but was associated not with the supernatant but with whole cultures. Interestingly, the phenotype was not detectable before a 2 day incubation period at 4°C. In *E. coli* the recombinant protein was not secreted but remained cytosolic. Purified protein was unable to aggregate *C. elegans*, suggesting it requires association with a second factor.

Western blot analysis was used to see if purified recombinant Mns was able to form any association with *C. elegans*. Nematodes were incubated with purified Mns for 15min and washed three times with PBS. The nematodes were suspended in SDS buffer and loaded for SDS-PAGE before blotting onto nitrocellulose. Antibodies raised against an Mns peptide detected the protein in the nematode preparation, suggesting that Mns was able to bind to the nematodes (**Fig. 2.7 (B)**). However, a no-nematode control also gave a positive signal, suggesting the protein was able to bind to the inside of the microcentrifuge tube and was not removed when washed with PBS, but was released by denaturation at 100°C.

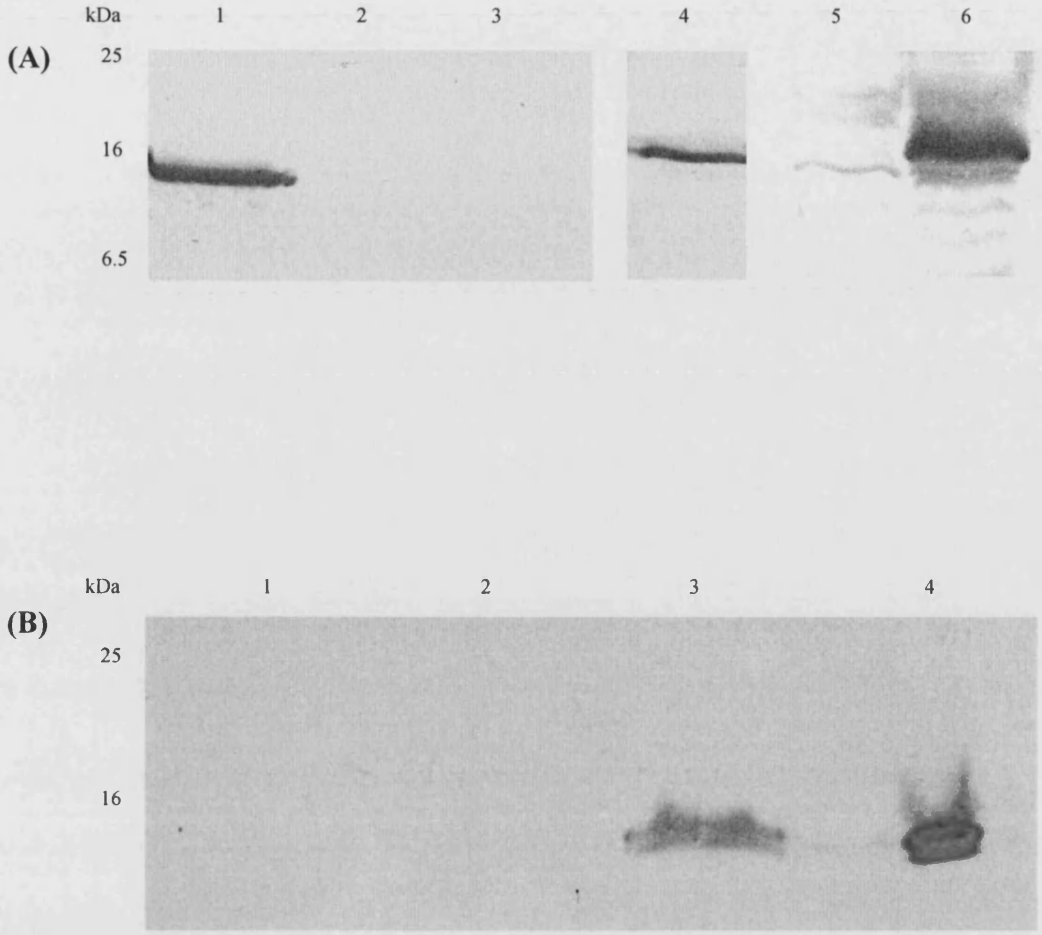


Fig. 2.7 Presence of Mns in aggregated nematodes. **(A)** Western blot investigating the presence of Mns in a sample of nematodes aggregated by *P. asymbiotica* supernatant. Lane 1, supernatant of *P. asymbiotica* grown at 30°C; 2, supernatant of *P. asymbiotica* grown at 37°C; 3, untreated *C. elegans*; 4, *C. elegans* clumped with 30°C *P. asymbiotica* supernatant; 5, *P. luminescens* TT01 supernatant Mns-knockout; 6, wild-type *P. luminescens* TT01 supernatant. **(B)** Western blot of Mns association with *C. elegans* and Kingscliff nematodes. *C. elegans* nematodes were incubated with recombinant Mns then washed in PBS. Anti-Mns detected the presence of Mns in the nematode preparation, as well as in a no-nematode control, indicating the protein was able to bind to a microcentrifuge tube. Mns was not detected in Kingscliff nematodes. Lane 1, Kingscliff nematode; 2, *C. elegans* incubated in PBS; 3, *C. elegans* incubated with Mns; 4, Mns incubated in microcentrifuge tube with no nematodes.

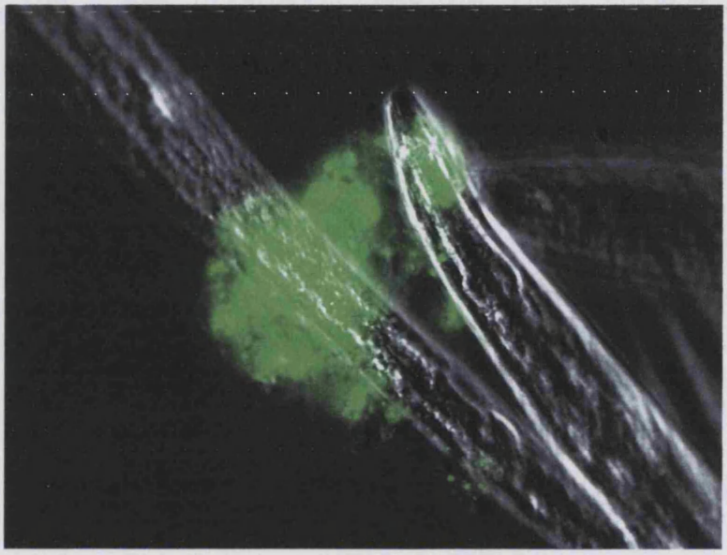


Fig. 2.8 *In situ* histochemical detection of Mns in a nematode aggregation. Confocal microscopy image of *C. elegans* aggregated with *P. asymbiotica* supernatant. The image shows the presence of Mns in the aggregation, revealed by green fluorescence from a fluorescent antibody against the anti-Mns peptide rabbit antibody.

2.3.6 Microscopy of Nematode Aggregates

E. coli cells induced to express Mns were examined by TEM and seen to form cell chains apparently due to an internal scaffold (**Fig. 2.9 (A)**). Furthermore, SEM showed that the cells were surrounded by a smooth cement-like matrix that binds them together. This matrix was seen to extend into the mouths of nematodes feeding on the *E. coli* (**Fig. 2.9 (B)**). This raises the possibility that worms are physically clumped by eating two ends of cell chains, or by physical entanglement in the *E. coli* filaments.

2.3.7 Mns Expression in *Photorhabdus* Cultures

P. asymbiotica cultures were grown at 28°C and 37°C and samples were taken at regular intervals in the first 24h, and less frequently up to 6 days. Expression of Mns from the cells *in vitro* was monitored by SDS-PAGE (**Fig. 2.10**). Mns was first detectable after 5h30. Due to a low concentration of cells in earlier time points we cannot exclude the possibility that Mns was secreted in the lag phase.

2.3.8 Mns Expression During Nematode Association and Insect Infection

To investigate possible Mns expression in the host nematode and during infection of insects, Western blot analyses were performed. Since no host nematode has been isolated for the sequenced *P. asymbiotica* strain ATTC43949, Kingscliff strain *P. asymbiotica* were used, with Kingscliff strain nematodes (Gerrard *et al.*, 2006). For infection, the model insect *G. mellonella* larvae were used.

Firstly, although no secretion of Mns was detected directly from Kingscliff *P. asymbiotica* grown in LB, a faint signal was seen from Kingscliff nematodes housing the bacteria after passage through an insect (**Fig. 2.11**). This suggests that Mns is produced during the symbiotic association between these organisms. Secondly, Western analysis showed that the protein is produced during infection (**Fig. 2.12**). No signal was detected in the first day of infection by Kingscliff *P. asymbiotica*, but a strong band was seen from day 2 onwards, suggesting the protein may be required throughout infection and persists after the insect is dead. The assay produced similar results with a second strain of *P. asymbiotica* and *P. luminescens* TT01, although expression of Mns was detected from day 1 in these infections. It should be stressed that although we did not see Mns after 24h infection by Kingscliff *P. asymbiotica*, it is

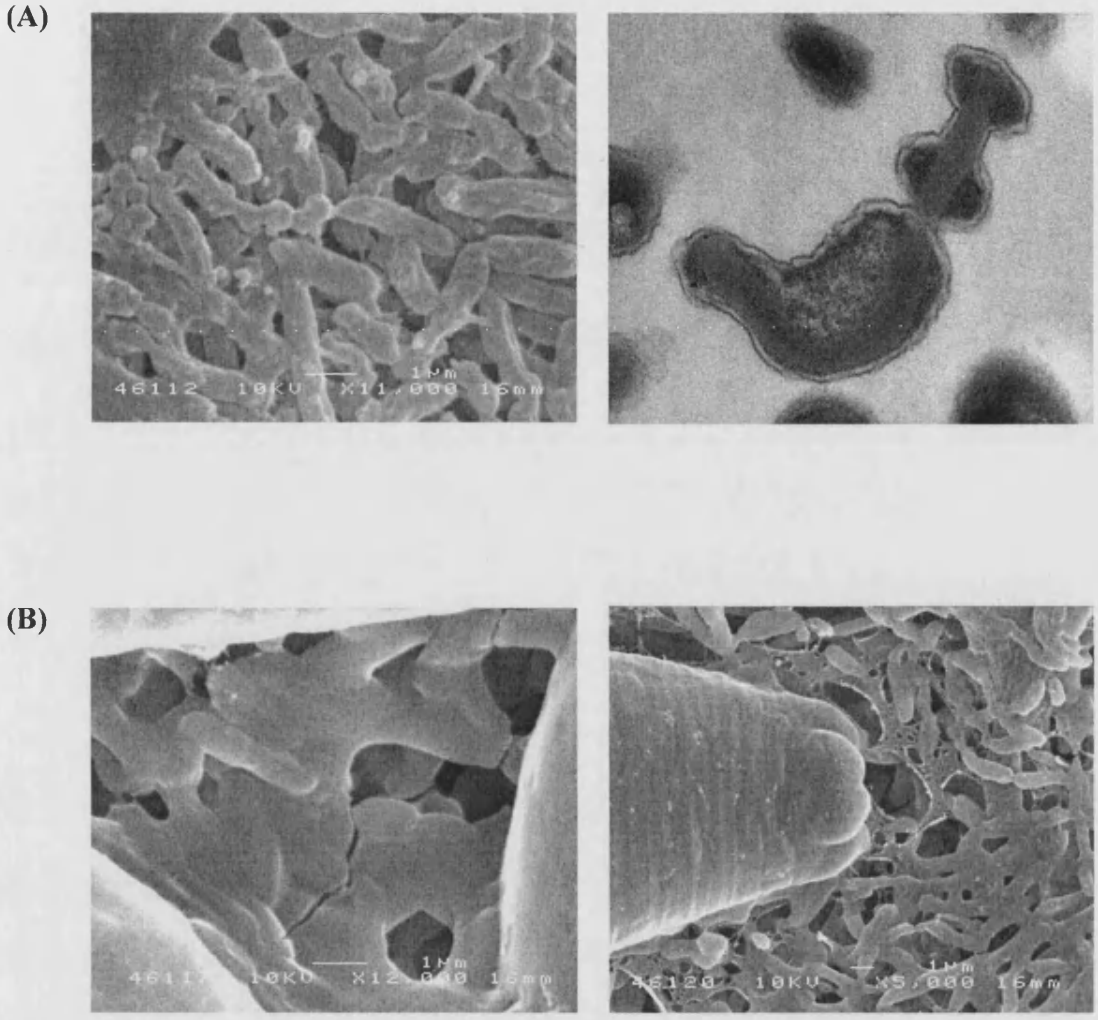


Fig. 2.9 Electron microscopy of Mns-expressing *E. coli*. SEM of Mns-expressing *E. coli* shows the formation of cells chains **(A,** left). Sections prepared for TEM show the presence of an internal scaffold that extends throughout the cells (right), and was not seen in pBAD30 empty-vector controls (not shown). Further magnification of the cell chains reveals a cement-like matrix that is continuous with the *E. coli* **(B,** left). This matrix extends into the mouth of feeding *C. elegans* (right).

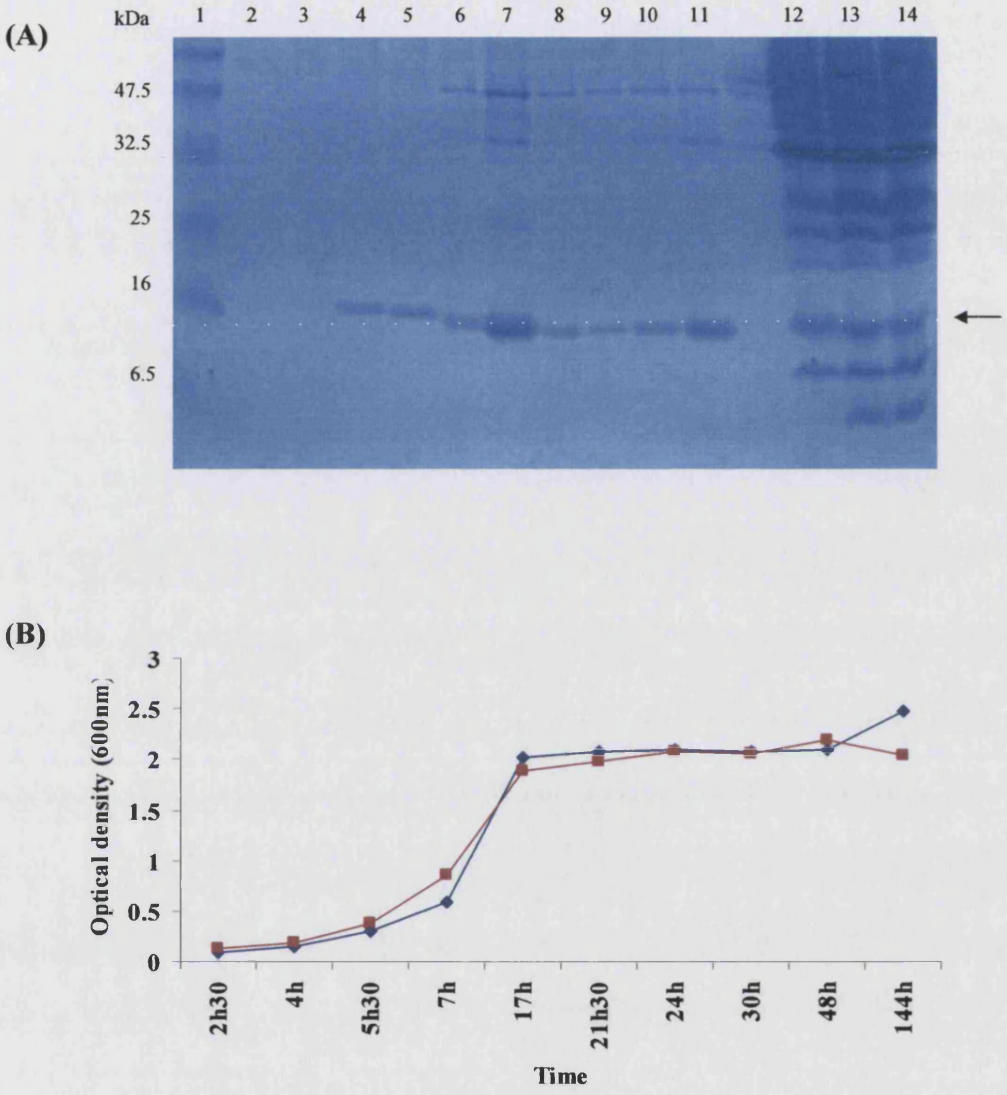


Fig. 2.10 Mns expression in *P. asymbiotica*. Mns expression in liquid culture was monitored from inoculation to 6 days at 28°C and 37°C. **(A)** SDS-polyacrylamide gel of *P. asymbiotica* culture grown at 28°C shows Mns (arrow) was first detectable 5h30min after inoculation (lane 4) and is present after 6 days (lane 11). Lane 1, protein marker; 2, 2h30; 3, 4h; 4, 5h30; 5, 7h; 6, 17h; 7, 21h30; 8, 24h; 9, 30h; 10, 48h; 11, 144h. (Lanes 2-11 show culture supernatant TCA preparations. Lanes 12-14 show cell preparations). Lane 12, 24h; 13, 48h; 14, 144h. **(B)** Optical density measurements of the growing culture (blue, 28°C; red, 37°C).

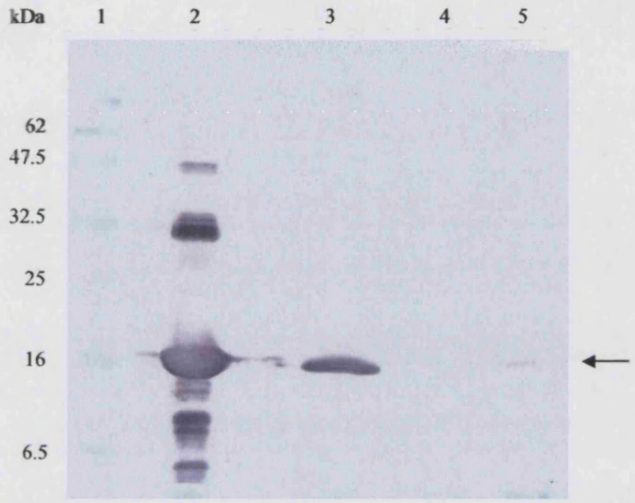


Fig. 2.11 Detection of Mns in nematode preparation. Western blot shows detection of Mns in a preparation of *P. asymbiotica* ATTC43949 supernatant, but not in a preparation from Kingscliff *P. asymbiotica* (lane 4). A very faint signal (arrow) was detected from crushed Kingscliff nematodes, suggesting there may be a low level of expression in the nematode. Lane 1, protein marker; 2, recombinant Mns; 3, *P. asymbiotica* ATTC43949 supernatant; 4, Kingscliff *P. asymbiotica* supernatant; 5, Kingscliff nematode.

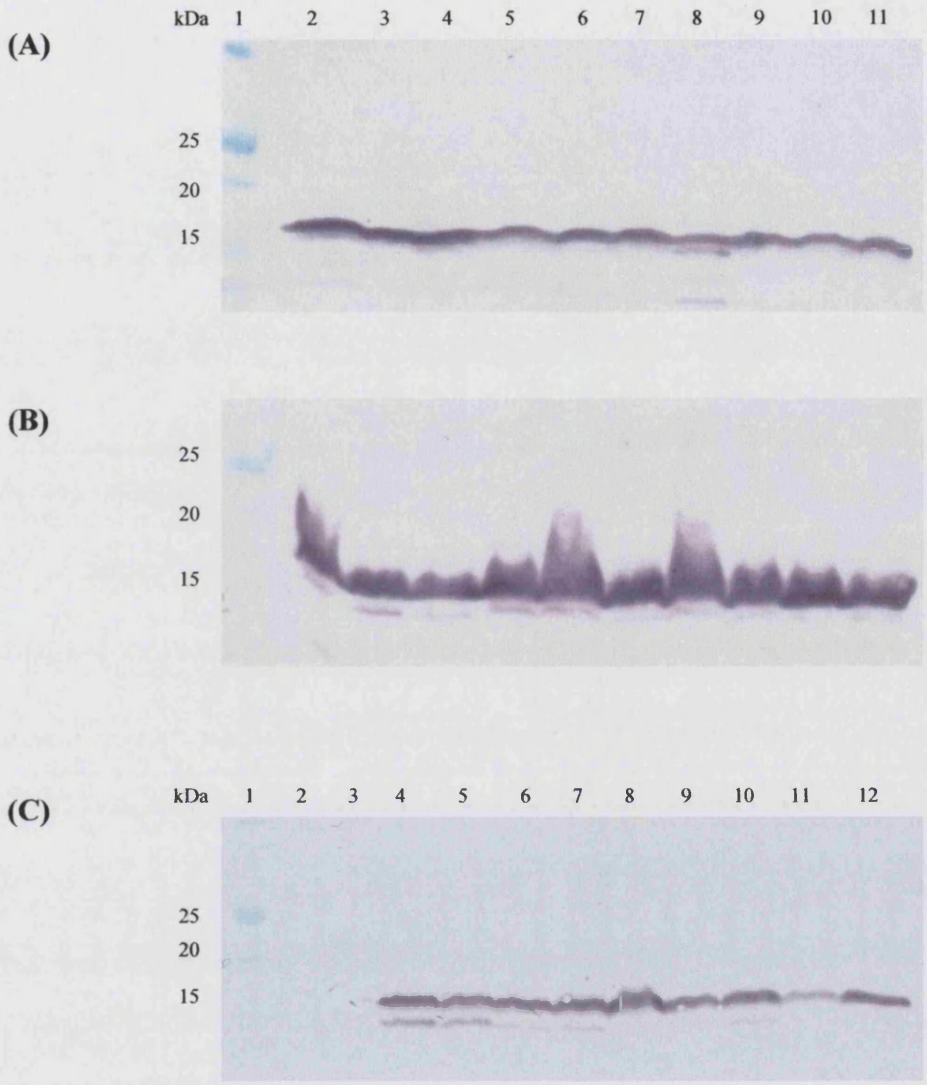


Fig. 2.12 Detection of Mns in infected *G. mellonella*. Insects were injected with *P. asymbiotica* Kingscliff strain and an infected insect was crushed on successive days for Western analysis. Mns was first detected from the second day of infection with three different strains of *Photorhabdus*. **(A)** *P. luminescens* TT01 infection. Lane 1, protein marker; 2, day 1 post infection; 3-10, days 2-10 post infection; 11, day 13. **(B)** *P. asymbiotica* ATCC43949 infection. Lane order as in **(A)**. **(C)** Kingscliff *P. asymbiotica* infection. Lane 1, protein marker; 2, *G. mellonella* uninfected; 3-7, days 1-5 post infection; 8, day 7; 9, day 8; 10, day 13; 11, day 15.

likely that the cell count inside the insect would be low at this time and expression may have been difficult to detect. Lastly, RT-PCR of the *mns* gene showed that the protein is constitutively transcribed and differences in expression levels must be translationally regulated (N.R Waterfield, unpublished data).

To examine possible differences in expression in different insect body tissues, *M. sexta* larvae were injected with *P. luminescens* and *P. asymbiotica.* After 24h the haemocytes, fat body and midgut were isolated from infected larvae and prepared for Western blot analysis. Anti-Mns monoclonal antibodies failed to detect Mns in any of the isolated tissues (not shown). This suggests that Mns is not expressed early in *M. sexta* infection.

2.3.9 Mns Toxicity

A BLAST search (National Centre for Biotechnology Information, www.ncbi.nlm.nih.gov/BLAST) using the predicted amino acid sequence of Mns shows the most closely related protein identified to be a binary insecticidal protein from *Bacillus thuringiensis* (data not shown). The high level of production and secretion of Mns also favour the possibility that Mns may be an insecticidal protein. To examine this hypothesis, recombinant Mns was administered orally and by injection to insect larvae.

2.3.9.1 Injectable Toxicity

To test the potential toxicity of Mns, a purified sample of recombinant protein was injected into *G. mellonella.* As a negative control, the buffer used to elute Mns from the anion exchange column was used, and showed no effect on the insects. After 24h, none of the injected insects showed any effects.

2.3.9.2 Oral Toxicity

First-instar *M. sexta* larvae were fed on an artificial wheat-germ diet supplemented with Mns, either purified or in the supernatant of lysed cells expressing the protein, and the weight of insects was weighed after 7 days. There was no significant difference in the weight of the insects compared to controls, indicating that recombinant Mns, purified or in *E. coli* cell lysate, is not orally toxic (**Fig. 2.13**).

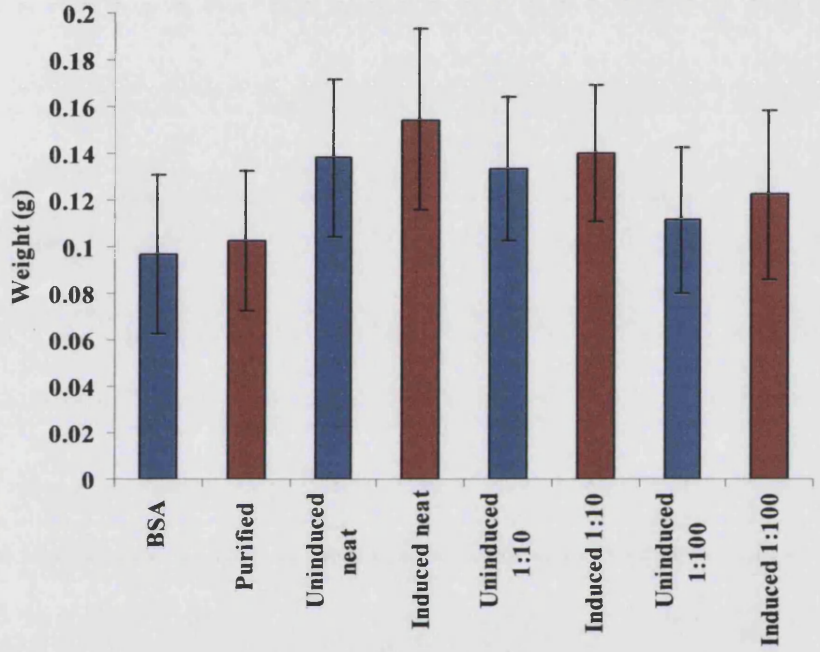


Fig. 2.13 Oral toxicity of Mns to *M. sexta* larvae. First instar *M. sexta* larvae were grown on artificial wheat germ diet supplemented with Mns and weight was measured after 7 days. There was no significant weight loss due to purified Mns compared to a bovine serum albumen (BSA) control. Similarly, induced Mns-pBAD30 *E. coli* and uninduced (pBAD30) cultures, lysed and diluted before application to the food blocks, showed no significant differences, suggesting that Mns is not orally toxic to these insects. Standard deviation bars are shown.

2.3.10 Biofilm Formation on Polypropylene Plates

Biofilm formation was quantified by growth of *P. luminescens* wild-type TT01 and Mns-knockout TT01 in the wells of a polypropylene plate. After 48h growth in static conditions, loosely bound cells were washed away and bound cells were stained with CV. Absorbance measurements of these suspensions showed that the Mns-knockout strain produces significantly less biofilm (**Fig. 2.14**). Similarly, growth of *P. asymbiotica* in static conditions at both 28°C and 37°C showed considerably less biofilm formation in the higher temperature, a condition that does not permit Mns secretion.

2.3.11 Pellicle Formation Assay

Biofilm pellicle formation was monitored in *P. luminescens* TT01 wild-type and Mns-knockout strains over 9 days. After 72h a thin broken ring of cells was first observed inside the glass vials at the air-liquid interface. The wild-type culture became darker in colour at the surface and produced a diffuse, grainy pellicle. At 72h the Mns-knockout culture was uniform in colour but produced a similarly grainy pellicle. By 96h, the ring of cells was thicker in both cultures but the Mns-knockout strain showed less attachment and a more heavily grained pellicle.

After 120h a uniform pellicle completely covered the surface of the wild-type strain culture (**Fig. 2.15**). This became increasing thick and porous at 144h, 168h and 192h. The Mns-knockout strain pellicle was granular at 120h and not uniform, but became more smeared in appearance by 144h and by 192h was thick and porous, and indistinguishable from that of the wild-type culture.

2.3.12 Transmission Electron Microscopy of *Photorhabdus* Colonies

To identify potential differences in biofilm structure between wild-type and Mns-knockout *P. luminescens*, colonies were grown on LB agar and sections were prepared for TEM. Furthermore, sections were incubated with anti-Mns antibodies and gold-conjugated anti-rabbit secondary antibodies to determine where in the colony structure Mns was located.

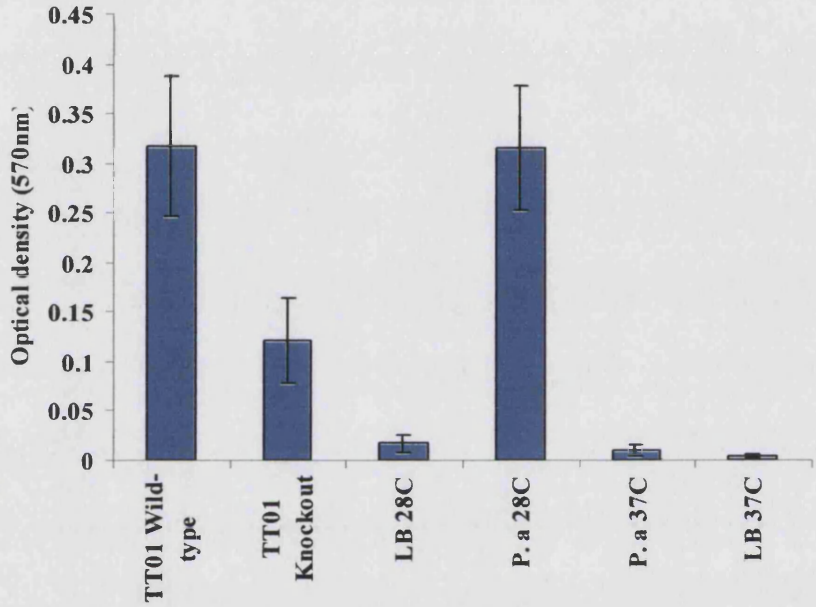


Fig. 2.14 Cell attachment assay. Biofilm formation was measured after 48h growth on polypropylene plates. Wild-type and Mns-knockout TT01 were grown at 28°C, and *P. asymbiotica* (P. a) was grown at both 28°C and 37°C. Loosely-bound cells were washed with water and attached cells were stained with crystal violet, which was removed with acetone-ethanol. The optical density of this solution was measured at 570nm. Knocking out Mns reduces biofilm formation over 48h, and as does growth of *P. asymbiotica* at 37°C. Standard deviation bars are shown.

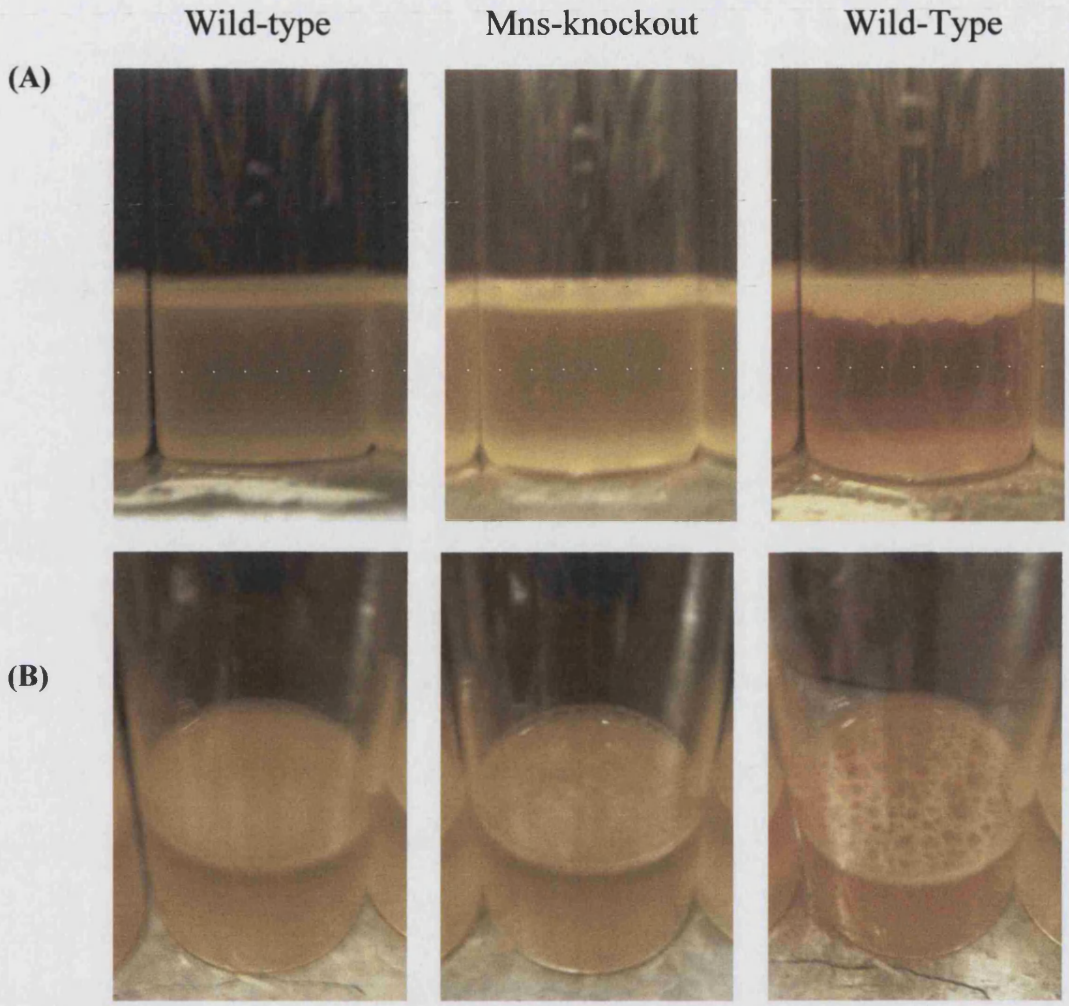


Fig. 2.15 Formation of *Photorhabdus* biofilm pellicles. Pellicle formation of *P. luminescens* was monitored over 9 days. Left panel shows wild-type TT01 and middle panel shows Mns-knockout. At 120h, wild-type TT01 produced a thin and uniform pellicle **(A).** Knocking out Mns resulted in a similar rate of pellicle growth, but the formation was marginally thicker at 120h and was not uniform but had a smeared granular appearance **(B)**. Both pellicles became progressively thicker and formed pores. Right panel shows pellicle of wild-type TT01 after 9 days (216h).

There were no discernible morphological differences between wild-type and Mns-knockout TT01 colonies. Cells in the colony were densely packed and separated in areas with large amounts of a highly convoluted extracellular material (**Fig. 2.16 & 2.17**). The presence of this material in both colonies suggests that its production is not dependent on Mns expression. However, immunogold labelling showed that Mns is associated with this material (**Fig. 2.18**). Mns was also present inside or on the surface of cells.

2.3.13 *GFP-Photorhabdus* Attachment to Glass Slides

Wild-type and Mns-knockout TT01 constitutively expressing GFP from the pHC60 plasmid were diluted from an overnight culture to an OD_{600} of 1.0 and dispensed into tissue culture plates containing glass coverslips. The cultures were grown without agitation and from individual wells the growing culture was periodically aspirated to monitor Mns expression by Western analysis. The glass slides were also removed to examine biofilm development.

Confocal microscopy showed a gradual increase in the number of cells attached to the glass slide. After 24h, Mns-knockout TT01 showed the development of microcolonies, evident as isolated masses of cells visible as regions fluorescing with GFP (**Fig. 2.19**). At the same time point, wild-type cells were attached as individual cells and the microcolonies were less mature and fewer in number. At subsequent time points there was little difference in the development of the biofilm between the wild-type and Mns-knockout cells.

P. asymbiotica slides, prepared in the same way but grown at 28°C and 37°C showed differences in both biofilm development and cell morphology. After 12h at 28°C, cells had begun to accumulate into microcolonies. These became more extensive by 24h, forming a network of cells on the slide (**Fig. 2.20**). Cells grown at 37°C were elongated and attached poorly to the glass slide. At 12h there were no microcolonies. After 24h, microcolonies were visible but contained fewer cells than those from the 28°C biofilm and the network of attached cells was less mature. Western blots showed that Mns was produced throughout the incubation period, detectable in *P. luminescens* supernatant and 28°C-incubated *P. asymbiotica* supernatant at the first time point, 6h (not shown).

(A)

(B)

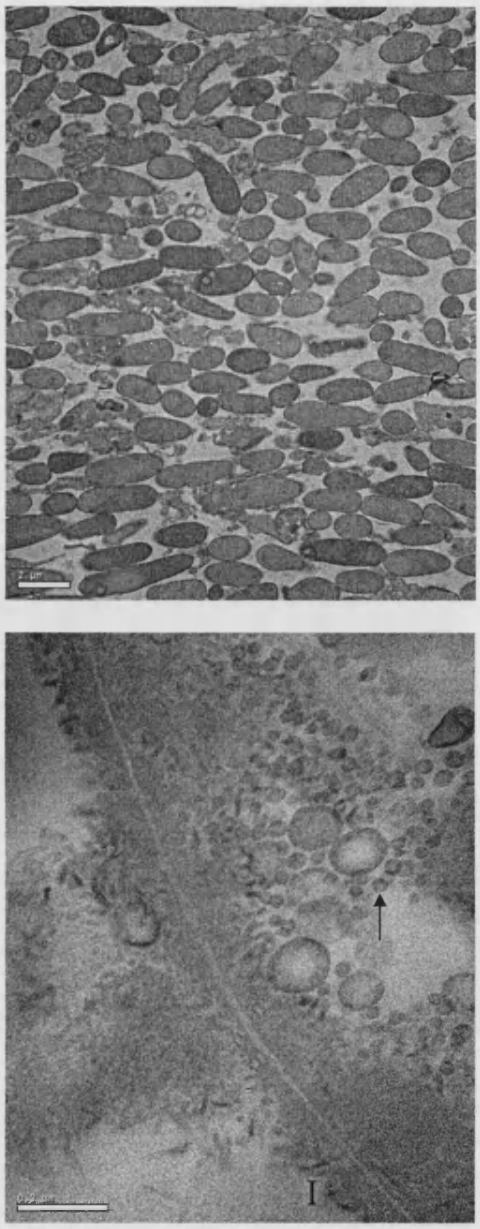
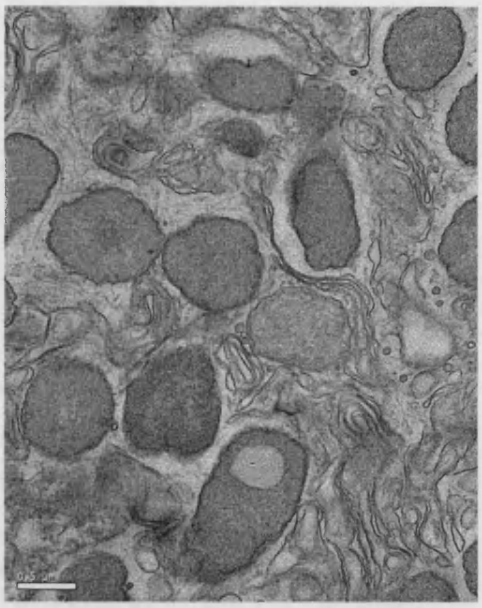


Fig. 2.16 Electron microscopy of wild-type biofilms. P. *luminescens* TT01 colonies were grown on agar and sections were viewed under transmission electron microscope. Cells appeared densely packed in the middle of the colony biofilm **(A)** and were surrounded by an extracellular matrix. Membrane vesicles (arrow) can be seen deposited close to the interface (I) with the agar surface **(B)**. Scale bar is 2μm **(A)** and 0.2μm **(B)**.

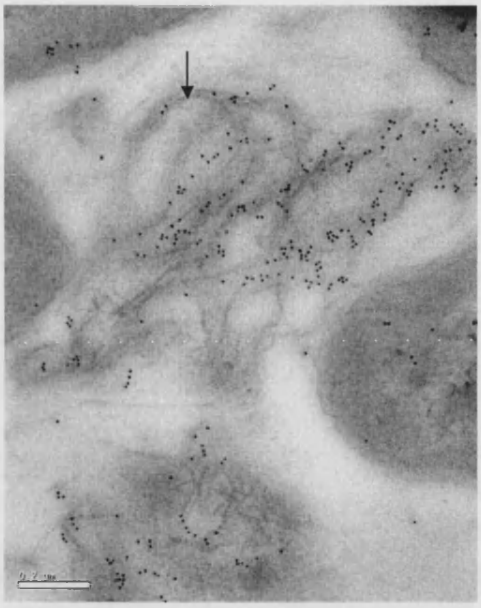
(A)



(B)

Fig. 2.17 Electron microscopy of wild-type and Mns-knockout *P. luminescens* colony biofilms. Sections of wild-type *P. luminescens* TT01 **(A)** and Mns-knockout TT01 **(B)** colonies were apparently identical when viewed by transmission electron microscopy. These sections show large amounts of extracellular matrix that surrounds the cells. Scale bar is 0.5μm **(A)** and 0.2μm **(B)**.

(A)



(B)

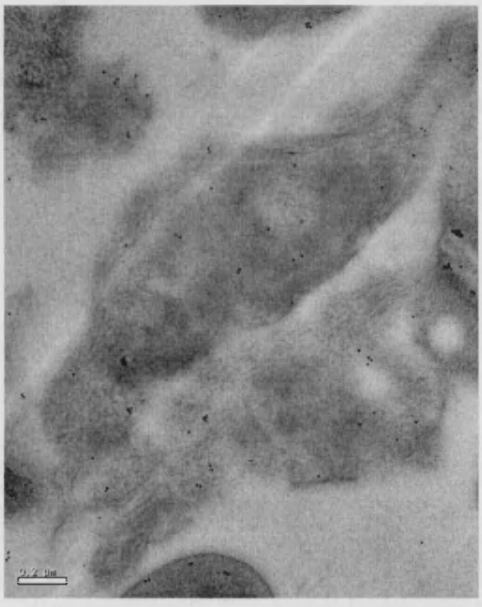


Fig. 2.18 Immunogold labelling of Mns in *P. luminescens* colony biofilm. The protein appears in the cell/at the cell surface, and decorates the extracellular material between cells of wild-type cells **(A)**. Mns-knockout *P. luminescens* **(B)** shows negligible labelling similar to that of no-primary antibody controls. Biofilms were grown on LB agar and sections were viewed by transmission electron microscopy. Gold conjugate on secondary (anti-rabbit) antibodies visible as electron dense dots (arrow). Scale bar is 0.2μm **(A)** and 0.2μm **(B)**.

(A)

(B)

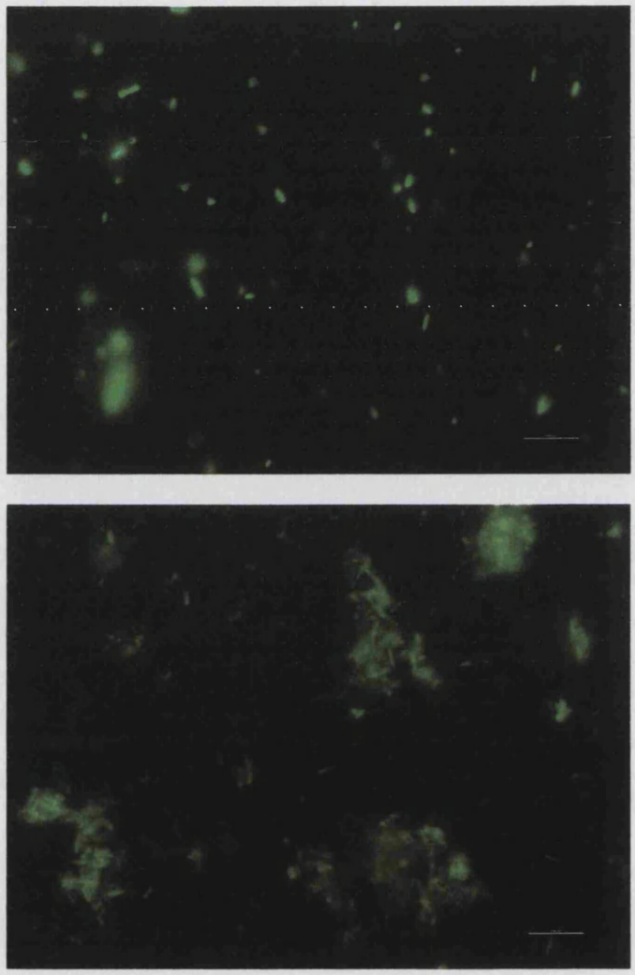


Fig. 2.19 Confocal microscopy of *P. luminescens* biofilms. Biofilm development was observed by growing cultures in static conditions in wells containing a glass coverslip. Cells that attached to the coverslip were visible through their expression of GFP. **(A)** Shows wild-type cells showed attached after 24h. The biofilm is poorly developed, with few microcolonies. **(B)** Mns-knockout strain TT01 viewed at the same time point showed greater biofilm formation, with more and larger microcolony formations. Bar is 1000μm.

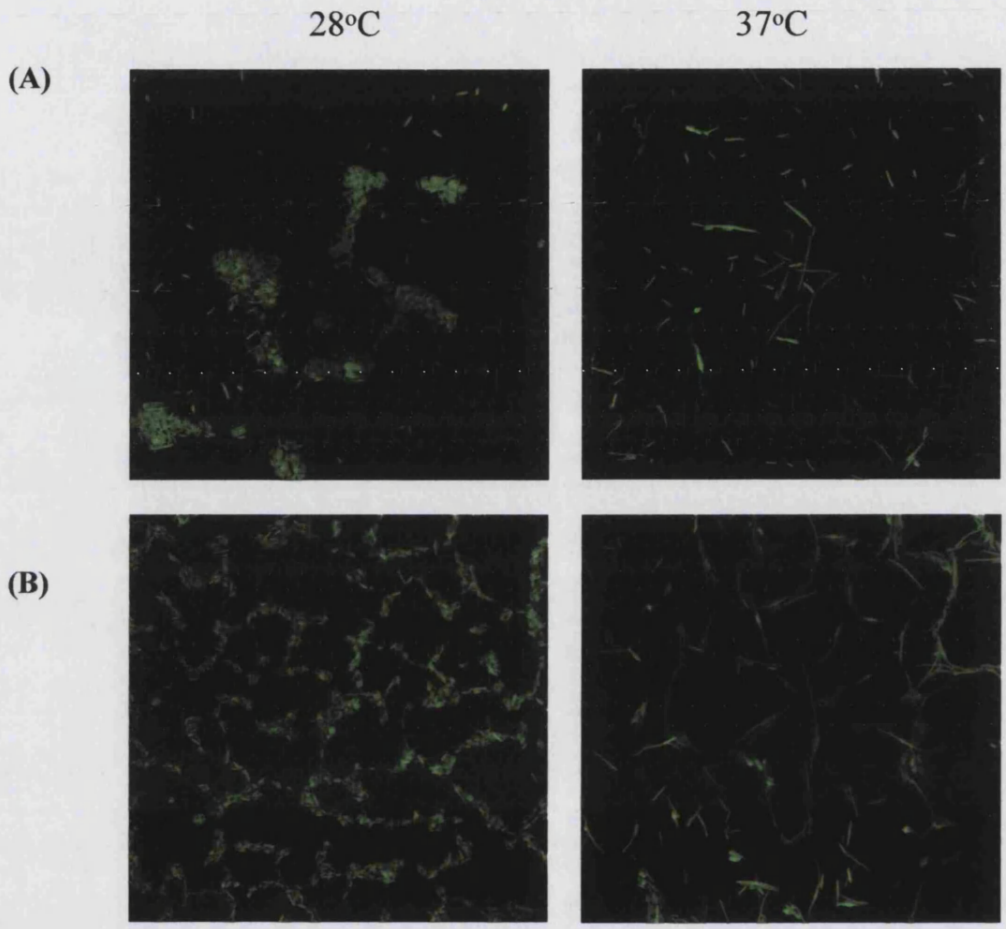


Fig. 2.20 Confocal microscopy of *P. asymbiotica* biofilms. Biofilms were observed with a confocal microscope after 12h (A) and 24h (B). Slides were prepared by growing a diluted culture of *P. asymbiotica* expressing GFP in static conditions at 28°C and 37°C in wells containing a glass cover slide. The slides were carefully washed to remove loosely bound cells, and attached cells were fixed before observation. Slides show differences in biofilm establishment, with the 28°C-grown cells forming microcolonies that extend into a network on the substrate. The cells grown at 37°C are considerably elongated by comparison to those grown at 28°C, and show much less adhesion at 12h. Although a loose network of cells forms, microcolonies contain much fewer cells.

2.3.14 Surface Plasmon Resonance

Surface plasmon resonance spectroscopy was used to determine the role of Mns in the initial adhesion of *Photorhabdus* to surfaces. Initially, bacterial cells washed in PBS were diluted in a series from 0.6E9 CFU/ml to 0.6 CFU/ml. 50µl of the suspensions were applied to an unmodified gold surface against a PBS standard, and measurements were taken in real-time over 1000s. As particles attached at the surface they caused changes in the angle of resonance. These changes were measured for both the bacterial sample and the PBS standard and the difference in the angle of resonance was plotted. These initial plots showed (a) that cells were attaching to the gold surface, and (b) allowed the optimal cell concentration to be identified because the angle of resonance was shown to be dependent on cell density. Above concentrations of 5.9E5 CFU/ml, the difference angle was in the order of 50m° (milli degrees) or greater after 1000s. A 1% crystal violet stain confirmed that cells had attached to the gold surface.

2.3.14.1 SPR Measurements of *P. luminescens* Cells

The above experiment was repeated at the higher cell densities using wild-type *P. luminescens* TT01 and the Mns-knockout strain. As the change in resonance angle was dependent upon cell density, the clearest differences in attachment to the gold surface between TT01 and Mns-knockout were seen at 5.9E8 CFU/ml. In these experiments, TT01 was injected onto the gold surface from channel 1 of the SPR instrument and Mns-knockout was applied from channel 2. Changes in the resonance angular were recorded for each channel and the difference angle has been plotted; an angle >0 indicated better adhesion of the knockout strain, and values <0 indicated better adhesion in presence of Mns. Over 1600s, Mns-knockout consistently showed better attachment than the wild-type strain (**Fig 2.21 (A)**). These data were supported by *E. coli* attachment. Cells induced to express recombinant Mns were less able to attach to the gold surface than control *E. coli* cells (**Fig. 2.21 (B)**).

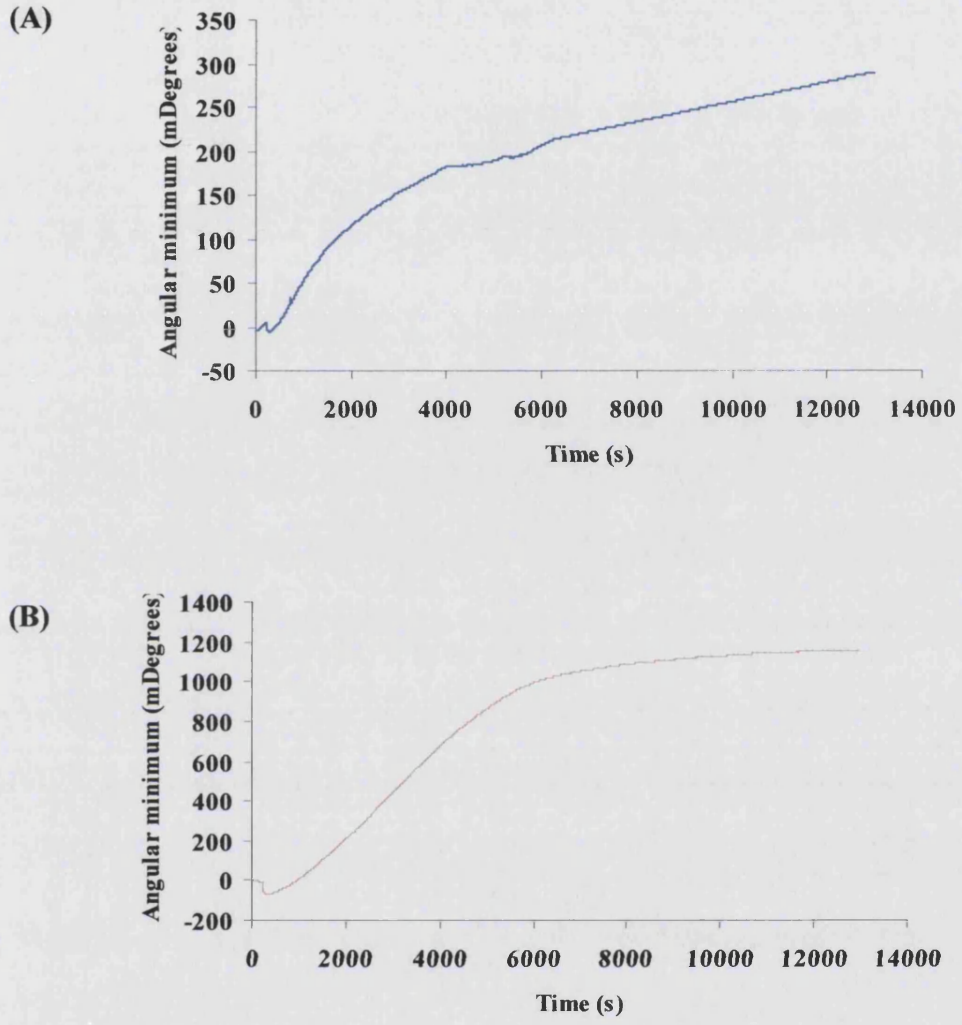


Fig. 2.21 Adherence of washed *P. luminescens* and *E. coli* cells measured with SPR. Cells were washed in PBS and applied to a gold disk. (A) Attachment of wild-type *P. luminescens* TT01 cells was monitored by SPR against Mns-knockout TT01. The increase in differential angle above 0° indicates greater attachment of the knockout strain compared to wild-type, suggesting the presence of Mns at the cell surface perturbs cell attachment. **(B)** Consistently, an *E. coli* control shows greater attachment than *E. coli* expressing recombinant Mns.

2.3.14.2 SPR Measurements of *P. luminescens* Cultures

When whole cultures (cells and their supernatants) were applied to the gold surface the differential angle showed an initial minor decrease, followed by an increase and stabilisation at ~100m° (**Fig. 2.22 (A)**); the wild-type strain initially showed greater attachment than the Mns-knockout strain, but after ~1000s Mns-knockout showed greater attachment and followed the attachment profile of the PBS-washed cells. Overall adhesion levels for both the wild-type and Mns-knockout cells were greater when cells were in their native supernatant than when washed and applied to the gold disk in PBS (data not shown).

2.3.14.3 SPR Measurements of *P. luminescens* Supernatants

To test if Mns in the supernatant caused a change in the SPR angle, wild-type and Mns-knockout TT01 were grown in LB to an OD_{600} of 1.6 and pelleted. The cell-free supernatant was applied to the gold surface and the resonance angle was measured over 4000s. As described above, an angle >0 indicated better adhesion in the absence of Mns. Here, attachment of extracellular material was detected from the supernatant of TT01 with a decrease in the SPR angle of 85m° when plotted against attachment of material from the knockout strain (**Fig. 2.22 (B)**). Anti-Mns antibodies were added to the gold surface and their attachment was monitored by a change in the SPR angle (not shown) indicating (a) that Mns is attaching to the gold surface, and (b) that the difference between the knockout and wild-type strains can be attributed to Mns.

2.3.14.4 Surface Modification

Modification of the attachment surface showed that *P. luminescens* are better able to adhere to more hydrophilic surfaces (**Fig. 2.23**). The lowest level of attachment was measured on the most hydrophobic surface tested, ODM. Oxidation of the ODM surface with ozone restored attachment measurements to the level of the unmodified gold surface. The Mns-knockout strain showed consistently greater levels of attachment than wild-type TT01 on all surfaces tested (data not shown).

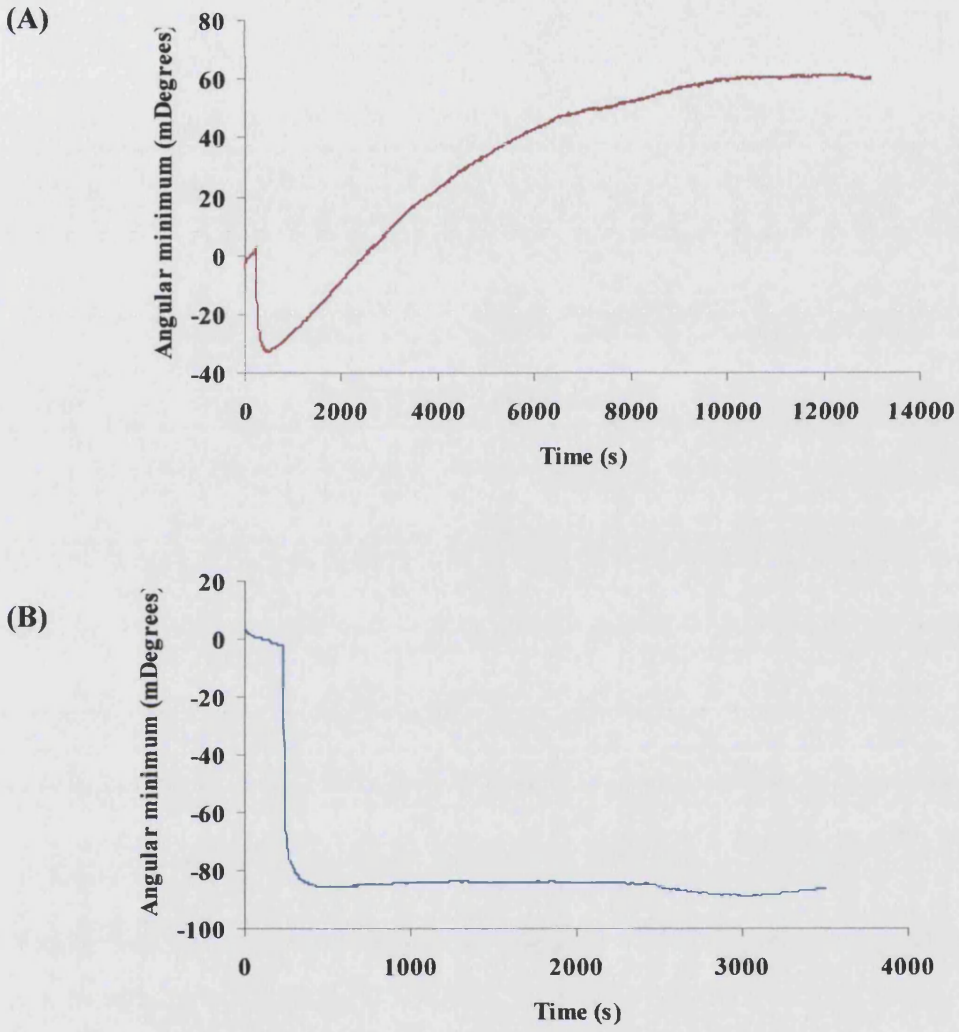


Fig. 2.22 Adherence of whole cultures of *P. luminescens* and material in the supernatant, measured with SPR. **(A)** Whole-culture preparations of wild-type *P. luminescens* TT01 and Mns-knockout TT01 were applied to a gold surface and monitored by SPR. The difference between the attachment of the two strains is shown, and indicates more adhesion by the wild-type strain in the first 500 seconds. Following this the differential angle increases, representing the adhesion of more Mns-knockout cells to the gold surface. **(B)** Plotting the attachment measurements of material in the supernatant of wild-type and Mns-knockout cells shows a decrease in the differential angle below 0°, indicating greater attachment from the wild-type strain and suggesting that Mns in the supernatant is binding to the gold surface.

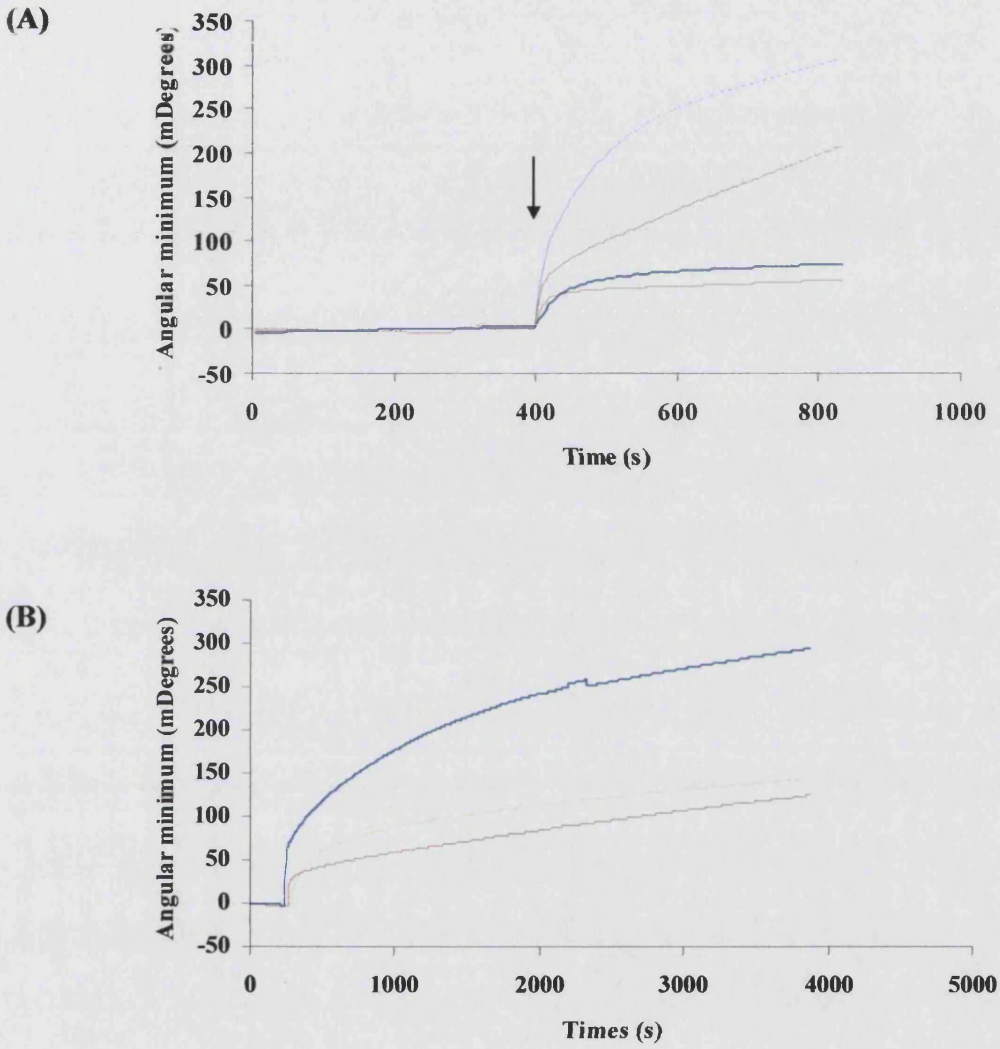


Fig. 2.23 Adherence of *P. luminescens* to Octadecylmercaptan-treated and unmodified gold surfaces, and measured with SPR. Attachment of *P. luminescens* Mns-Knockout TT01 to an unmodified gold surface (blue) and a gold surface treated with octadecylmercaptan (red). **(A)** Attachment was measured in real time with SPR, and showed greater adherence to the more-hydrophilic unmodified gold surface. Cells were diluted at 4.9E7 (lower lower traces) and 4.9E8 cfu/ml (upper traces) for each surface, and applied to the gold disk at 400s (arrow). Greater attachment was recorded from the higher cell densities as expected. **(B)** Adhesion of *P. luminescens* to a surface of intermediate hydrophobicity, mercaptopropionic acid (yellow), was stronger than on the most hydrophobic surface and weaker than on unmodified gold. Mns-knockout TT01 adhered better than wild-type TT01 on each surface tested (not shown).

2.3.15 Swimming and Swarming Assays

P. luminescens TT01 did not show any swarming activity. On swimming plates, both wild-type and Mns-knockout TT01 were capable of swimming motility, but there was no significant difference between the advance of either cell type across the media (not shown).

2.4 DISCUSSION

2.4.1 Nematode Interaction

An interaction exists between *P. asymbiotica* and *C. elegans* nematodes. Nematodes exposed to the culture supernatant of *P. asymbiotica* aggregate into small masses, usually tethered at the tail. Within this mass, material accumulates, either collected from the supernatant or secreted by the nematodes. This phenomenon was found not to occur with the supernatant of cultures grown at higher temperatures.

Two-dimensional gel electrophoresis has been performed on *P. asymbiotica* supernatants in an effort to identify factors expressed at 37°C and not at 28°C, and therefore may be involved in human infection. One protein, revealed by mass spectroscopy to be a novel protein identified as plu1537 by Turlin *et al.* (2006), was cloned into *E. coli* and appeared to replicate the aggregation phenotype. The protein, Mns, was poorly secreted from *E. coli*, and nematode aggregations were associated with aggregated cells. When these aggregations were examined by scanning electron microscopy, the cells appeared to be covered in a continuous cement-like matrix (**Fig. 2.9**).

Observation of *E. coli* cell aggregates by transmission electron microscopy revealed that the cells showed incomplete septation, and appeared to be sated with a large internal fibrous structure (**Fig. 2.9**). The extent of this cell aggregation phenotype associated with Mns overexpression was exaggerated by static incubation of cultures at 4°C, although it was detectable as autoaggregation even in shaking cultures (G. Yang, unpublished observation). Cellular aggregation has been reported previously and appears to be related to the formation of bacterial biofilms. For instance, in *Staphylococcus epidermidis*, cell aggregation is associated with the polysaccharide intercellular adhesin PIA, which is an important virulence factor because of its role in the establishment of microcolonies during biofilm formation on biomaterials (Mack *et al.*, 2000). PIA is a surface associated polysaccharide produced by proteins encoded at the *icaADBC* locus. Mutant strains unable to express these genes are biofilm negative. Similarly in *E. coli*, autoaggregation is mediated by Antigen 43, which also contributes to biofilm formation (Danese 2000).

2.4.2 Biofilm Formation

The ability of *P. asymbiotica* supernatants to aggregate *C. elegans*, and the observation of aggregation between Mns-expressing *E. coli*, encouraged us to study the potential of *Photorhabdus* to form biofilms, with a particular focus on the role of Mns in this process. The method of Ferrières and Clarke (2003) to quantify biofilm attachment has been used previously to study *Photorhabdus* (D.J. Clarke, personal communication). Here, dilute *P. luminescens* cultures were grown in static conditions and cell attachment to the abiotic surface was quantified by staining and absorbance measurement. The extent of biofilm formed by an Mns-knockout strain was significantly reduced in comparison to the wild-type control after 48h, suggesting that Mns promotes cell-cell or cell-substrate adhesion, either of which may enhance biofilm formation (Dunne, 2003). In the absence of Mns cells are less able to form these interactions.

2.4.2.1 P. asymbiotica Biofilms

Since *P. asymbiotica* is able to grow at 37°C as well as at lower temperatures, we measured biofilm formation in two temperature environments, and marked differences were observed. At 28°C, *P. asymbiotica* grew a biofilm very similar to that of wild-type *P. luminescens* at the same temperature. At 37°C, very little attachment to the abiotic surface was measured. It is important to note that although the secretion of Mns and few other proteins differ in these two conditions, it is expected that several other factors will also show differential regulation. For instance, cellulose is an important constituent of *Salmonella* biofilms, and its synthesis is dependent upon different environmental factors, including temperature (Solano *et al*, 2002). In *E. coli*, curli and cellulose synthesis are also regulated in a temperature-dependant manner, affecting biofilm formation and cell-cell aggregation, although the mechanisms that control these events downstream of the transcriptional regulator CsgD are different (Brombacher *et al.*, 2006); the reduction of biofilm formation by *P. asymbiotica* at 37°C is consistent with a lack of Mns, but absence of this protein is unlikely to be the only factor.

2.4.2.2 Pellicle Formation

The formation of a biofilm pellicle was monitored in *P. luminescens* over an extended period by growing dilute cultures of wild-type and Mns-knockout TT01 statically in glass vials. The biofilms were observed at 24h periods and the study revealed little overall difference after 9 days, with thick pellicle formations produced with a porous appearance. After 72h deposits were visible at the air-liquid interface of both cultures. However, this attachment was marginally less in the Mns-knockout cultures and by 120h the wild-type strain had formed a uniform pellicle, while Mns-knockout pellicles at the same stage were more granular in appearance.

This shows that Mns is not essential to the formation of the biofilm pellicle, and the differences between the Mns-knockout strain and wild-type TT01 were subtle in comparison to some EPS-mutant strains of *P. luminescens* (M.R. Amos, unpublished observations). However, the difference in the pellicle development observed after 72h is consistent with the difference measured in cell attachment to polypropylene wells after 2 days.

2.4.3 Confocal Microscopy of Biofilm Development

Biofilm development was monitored more closely with GFP-expressing *P. luminescens* attaching to glass slides, which could be viewed by confocal microscopy. As expected, individual cells initially attached to the surface, and formed microcolonies that gradually developed into more mature biofilm structures. Compared to wild-type cells, Mns-knockout TT01 showed greater attachment within the first 24h. Isolated masses of microcolonies were identified, and were both larger in size and more numerous on the glass. This finding contrasts with our studies examining attachment to polypropylene and pellicle development in glass vials, but supports our SPR data which finds greater attachment of Mns-knockout cells very early in biofilm establishment.

Biofilm formation by *P. asymbiotica* grown at 28°C and at 37°C was more consistent with the polypropylene-attachment study and pellicle formation assay. Cells grown at the lower temperature and shown to produce Mns formed microcolonies within 12h and by 24h the microcolonies had extended into a network on the slide. The 37°C-grown cells did not produce Mns and showed much less adhesion. The cells were also elongated, and formed a loose network of cells constructed from fewer and smaller colonies. Clearly, these findings further complicate the assessment of Mns, particularly as the microcolonies formed by Mns-knockout TT01 were not unlike those formed by *P. asymbiotica* at 28°C, suggesting that Mns is not critical to the process of cell migration to form microcolonies or cell-cell adhesion.

2.4.4 Transmission Electron Microscopy of Biofilms

The colony morphology of wild-type and Mns-knockout TT01 was examined by transmission electron microscopy. Differences between the two strains could not be discerned, with both colonies characterised by cell-dense assemblies with areas of convoluted extracellular material. The extracellular matrix produced by bacterial cells defines the architecture of a biofilm and the spatial arrangement of cells. Matrices are composed largely of extracellular polysaccharides, which maintain a highly hydrated layer that prevents desiccation and may promote bacterial adhesion (Sutherland, 2001). Other molecules may also exist in a matrix, such as proteins and nucleic acids, and discarded particulate structures such as pili and flagella (Branda *et al.*, 2005). Membrane vesicles are ubiquitous in Gram-negative biofilm matrices, and may contain virulence factors such as proteases, or bind to antibiotics. It has also been suggested that these vesicles could stabilise polymers or alter the substratum surface to facilitate adhesion (Schooling & Beveridge, 2006). Membrane vesicles were observed in both of the *P. luminescens* colonies examined here, particularly at the biofilm-substratum interface.

Wild-type TT01 extracellular matrix was decorated with gold particles when anti-Mns antibody was probed. Since this matrix was present and apparently very similar in the Mns-knockout strain, we conclude that Mns is not a major constituent of this material, but is in association with it. The location of Mns is consistent with a hypothesis in which the protein modifies surfaces, so altering the attachment of cells to the substrate or to one another. Mns did not appear to be deposited at the biofilm-substratum interface, but was present throughout the biofilm. Smaller amounts were also seen on or within cells.

Gold particles revealing the location of Mns were in many cases aligned along filamentous structures. It is tempting to compare these with the filamentous structures observed within the cells of *E. coli* induced to overexpress Mns, or assembling on the outside of *E. coli* cells grown under curli-proficient conditions (Bian & Normark, 1997). However, since these filaments were not absent from the matrix of the Mns-knockout biofilm, Mns must associate with them but not significantly alter their morphology.

2.4.5 Surface Plasmon Resonance

Surface plasmon resonance spectroscopy was used to examine the possible role of Mns in the initial stages of attachment. Although SPR has been extensively used to study the attachment of proteins and other polymers, few reports describe its use in the detection of bacterial cell attachment, despite being sensitive to such phenomena and being able to provide analyses in real time. Here, SPR has shown differential attachment of *Photorhabdus* cells to different surfaces, and in accordance with Jenkins *et al.* (2004), has been able to discriminate between the absorbance of wild-type and mutant cells.

SPR data for *Photorhabdus* showed essentially that Mns-knockout TT01 adhere to artificial surfaces better than wild-type cells. This perturbation of cellular adhesion is supported by data collected for *P. asymbiotica* grown at 28°C, which showed less adhesion to the gold substrate than cells grown at 37°C and therefore not expected to secrete Mns (data not shown). More directly, we expressed Mns as a recombinant protein in *E. coli* and found these cells were less able to attach than *E. coli* control cells.

Contrasting with these data, wild-type TT01 cell-free supernatant showed greater absorption than Mns-knockout supernatant. Since the only factor that differs between these strains is Mns, we conclude that Mns itself is adhering to the gold. Interestingly, when whole cultures of *Photorhabdus* were analysed, as opposed to washed cells resuspended in PBS, we observed an initial dip in the difference angle of resonance, which means the wild-type culture initially showed greater absorption than the knockout strain. This was followed by a large increase in the angle to above 0° after 2000 seconds. This pattern was not seen with SPR of the washed cell preparations. A possible interpretation of these observations is that secreted Mns in the wild-type supernatant initially binds the gold surface. The bound proteins modify the surface and affect subsequent adhesion of the wild-type cells. In the Mns-knockout whole culture there is no surface conditioning by Mns, and the angle of resonance reflects simply attachment of the cells to the gold substrate. Indirect support for this model comes from *P. asymbiotica* attachment, which was greater in the first 600 seconds after application to the gold surface from the culture grown at 30°C. As with wild-type *P. luminescens*, the possibility exists that this change in the angle of resonance is due to Mns adsorption to the substrate.

The ability of Mns to bind to abiotic surfaces is supported by Western blot analysis. When confirming the presence of Mns in aggregated nematodes, a control sample of Mns in a microcentrifuge tube was prepared without nematodes. The tube was thoroughly washed and SDS-gel loading buffer was boiled inside the tube. Western analysis of this sample showed Mns had been retained in the tube prior to denaturation. Furthermore, SPR showed that Mns had bound to gold disks that had been incubated with *P. luminescens* supernatant; anti-Mns antibodies were applied to the disk and preferentially bound to the gold surface where proteins of the wild-type supernatant had been allowed to absorb.

2.4.5.1 Surface Hydrophobicity

We observed in this study that *Photorhabdus* preferentially adhered to more hydrophilic than hydrophobic surfaces. The initial interactions between bacterial surfaces and a substrate surface may be attractive or repulsive, depending on the surface chemistries, but as a general rule bacteria with hydrophobic cell walls are considered to prefer hydrophobic surfaces and those with hydrophilic cell walls prefer hydrophilic surfaces (Katsikogianni & Missirlis, 2004). The surface hydrophilicity of *Photorhabdus* has not been measured, and this parameter is known to vary between different species and different strains.

Contrasting with the data presented for *Photorhabdus*, a greater occurrence of adhesion to hydrophobic surfaces than to hydrophilic surfaces has been reported for the *S. epidermidis* (Cerca *et al.*, 2005). Interestingly, for this species Cerca *et al.* (2005) found that the hydrophobicity of the bacterial cell had little or no bearing on adhesion. Furthermore, there was no direct relationship between the amount of biofilm formed and the extent of initial adhesion, indicating that high levels of initial adherence need not necessarily lead to extensive biofilm formation. We observed the same phenomenon with the Mns-knockout mutant of *Photorhabdus*. Wild-type cells showed less attachment in the very early stages of biofim formation, as measured by SPR, and more dense biofilm formation after 48 and 72h, revealed by crystal violet staining assays of biofilms on polystyrene plates and the extent of pellicle formation.

For a hydrophilic strain of *E. coli* K12, Ploux *et al.* (2007) showed that surface hydrophobicity made little difference to cell adhesion, but the spatial distribution of attached bacteria and the kinetics of biofilm formation were clearly affected. On the hydrophobic surface bacteria were organised into clusters, while a ramified arrangement was seen on the hydrophilic self-assembled monolayer. These morphologies showed similarities to those of polymer thin liquid films on the same surfaces, suggesting differences in substrate chemistry, rather than biological processes, govern structural arrangements (Ploux *et al.*, 2007). Clearly, these structures may respond differently to shear stresses, which will affect the development and maturation of the biofilm. The different physiological conditions provided by these structures will also lead to different bacterial responses, e.g. in the kinetics of bacterial detachment.

The role of bacterial surface structures should not be ignored. Several studies have implicated these factors, particularly in events following initial attachment. Using surface plasmon resonance, Jenkins *et al.* (2005) have shown a role of type IV pili early in the attachment of *P. aeruginosa*, and type I fimbriae are required for attachment of *E. coli* after contact is made to several abiotic surfaces (Pratt & Kolter, 1998). The presence of surface structures and complexity of bacterial surface polymers, as well as changes in the expression of these factors in response to environmental change, contribute to variability in the attachment of bacteria to substrates. These factors affect adhesion so much that it is almost impossible to make accurate predictions on physicochemical models (Katsikogianni & Missirlis, 2004).

2.5.4.2 Surface Conditioning

Surface conditioning occurs when a foreign surface is modified by the adsorption of water, lipids, proteins etc, which may change the potential for bacterial interaction. The presence of conditioning films on surfaces further complicates attachment studies. For instance, serum and tissue proteins such as fibrinogen may promote or suppress bacterial adhesion, and may be present in the aqueous environment or bound at the bacterial or substrate surface (Katsikogianni & Missirlis, 2004). Patel *et al.* (2007) reported that in the presence of absorbed serum proteins initial adhesion of *S. epidermidis* was considerably suppressed. Interestingly, after 48h inter-bacterial adhesion overcame this effect and lead to the formation of a robust mature biofilm. By contrast, the presence of absorbed fibrinogen on self-assembled monolayers allows for greater attachment of *S. aureus*, which has several cell wall-associated fibrinogen binding proteins (Tegoulia & Cooper, 2002).

Conditioning films can alter surface roughness and converge the hydrophobicities of hydrophobic and hydrophilic substrata, enhancing the ability of hydrophilic bacteria to adhere to a hydrophobic surface (Bakker *et al.*, 2004). In the present study, Mns appears to be a general adhesive factor that is highly secreted from *Photorhabdus*. We propose that Mns in the bacterial supernatant can bind to substrates and modify adhesion of the bacteria. Since the effect is not a complete gain or loss of attachment, the protein may allow some plasticity in colonisation during the infection process, e.g. cells transiently attaching at the trachea. Finally, it should be stressed that attachment to abiotic surfaces is not the same as attachment to living or devitalised tissue. The

modification by Mns of adhesion on gold and modified gold surfaces may not necessarily translate to the same modifications inside an insect or nematode host.

2.4.6 Swimming and Swarming

In motility assays on semisolid agar, we found that *P. luminescens* were able to swim but were not motile on the more solid swarming plates. Swimming is usually associated with a single polar flagellum and allows movement through liquids, while swarming is promoted by the production of numerous peritrichous flagella that enable bacteria to move over surfaces or through more viscous media (Murray & Kazmierczak, 2006).

We found no difference in swimming motility between wild-type cells and those unable to express Mns; the Mns-knockout mutation does not interfere with this form of motility. However, in the presence of specific carbon and nitrogen sources, it may be possible to induce swarming motility in *Photorhabdus*, as has been reported for *Pseudomonas* (Köhler *et al.*, 2000). This may be worth further investigation because of the differences observed between Mns-knockout and wild-type biofilms. Like biofilm formation, the swarming state involves multicellular behaviour, with apparent organisation and the production of an extracellular slime. A correlation between the ability to form a biofilm and the ability to swarm on semisolid media has been described previously for *P. aeruginosa*, in which swarming requires flagella, type IV pili, and the production of rhamnolipids. Rhamnolipids are amphipathic molecules, composed of a hydrophobic lipid and a hydrophilic sugar moiety. These biological surfactants enhance cell surface hydrophobicity and are able to reduce the friction between cells and surfaces, a property that implements them in maintaining fluid channels in mature biofilms and allowing flagella-mediated propulsion when swarming (Köhler *et al.*, 2000).

Biosurfactants are produced by a wide variety of microorganisms and have very different surface properties. High molecular weight biosurfactants may be composed of polysaccharides, lipopolysaccharides, proteins, or lipoproteins (Ron & Rosenberg, 2001). Complex mixtures of these components are also found, with the presence of the attached groups providing surface activity to the protein. Alternatively, as in the hydrophobin HFBII, it is the properties of the exposed amino acid sidechains that confer an amphipathic nature to the molecule. HFBII has a largely hydrophilic surface but a distinct patch of aliphatic residues that create a hydrophobic patch. The amphiphilicity of the molecule drives it to form dimers at low concentrations or assemble into monolayers at interfaces (Hakanpaa *et al.*, 2004). Hydrophobins are the most surface-active proteins known, are unique to filamentous fungi and are characterised by eight conserved cysteine residues. Like HFBII, SC3 from *Schizophyllum commune* is a moderately hydrophobic protein known to self assemble into an insoluble amphipathic film. On the hydrophobic side of the film a mosaic of parallel amyloid-like rodlets form, and the assembly process only occurs after interaction with a hydrophobic-hydrophilic interface (de Vocht *et al.*, 2000).

Mns has only one cysteine residue and as such is unable to form the disulphide links that stabilise the hydrophobins and prevent their aggregation inside the cell or in the aqueous environment (de Vocht *et al.*, 2000). However, we have found that Mns aggregates and the apparent ability to bind to artificial surfaces suggests the protein has surfactant-like qualities. Furthermore, Mns modifies the attachment of cells to surfaces and affects the formation of biofilms, consistent with a role of the protein at interfaces.

Hydropathy profiles of proteins may be used to identify potential transmembrane domains and clusters of hydrophobic residues, but such plots cannot be used to determine if these residues are buried inside a globular protein or if they are exposed and produce a state of amphipathicity. Without detailed structural data it remains unclear whether or not Mns has the structural properties of a surfactant. The possibility that Mns is glycosylated should not be ignored. Glycoproteins have been found increasingly in insect and mammalian pathogenic bacteria, and appear to be surface associated or secreted, with suggested roles in cell adhesion and interactions with their hosts (Benz & Schmidt, 2002).

Glycosylation has the potential to alter the surface properties of a protein, and conceivably produce amphipathic characteristics. No consensus sequence has been identified for the covalent linkage of carbohydrate groups in prokaryotes, although the N-linkage environment in eukaryotic proteins of Ser/Thr-Xaa-Asn (in which Xaa may be any amino acid except proline) has been found in some cases (Benz & Schmidt, 2002). Interestingly, Mns contains two such sites (Asn positions 21 and 50, identified with the NetNGly 1.0 server, Technical University of Denmark) **(Fig. 2.24)**.

Mns is predicted to be 14.9kDa and migrates very close to a 15kDa marker when loaded on SDS-polyacrylamide gels. Any bound carbohydrates must be small enough to not produce an aberrant migration, but carbohydrate residues have been observed elsewhere to account for just 4% of the total glycoprotein weight (Benz & Schmidt, 2002). Such an increase is unlikely to be detected by SDS-PAGE without the parallel examination of glycosylated and de-glycosylated protein. Such analysis deserves attention.

A lack of glycosylation in recombinant Mns is a possible explanation for our inability to efficiently reproduce nematode aggregation with the recombinant protein. Alternatively this may stem from the absence of a secondary factor that was not available in *E. coli* or to the purified protein. Unfortunately, due to protein aggregation and apparent adhesion to filter concentrators, we were unable to purify Mns directly from *Photorhabdus*, despite it comprising approximately 30% of all secreted protein and being relatively pure in the supernatant. Lastly, our recent findings suggest that the protein, rather than being glycosylated, may be able to bind to sugars (data not shown).

2.4.7 Mns Toxicity

No injectable or oral toxicity of Mns towards insects was observed in the present study, but there still remains a possibility that the protein is a secreted toxin. As mentioned above, when produced heterologously Mns may lack post-translational modification or a co-factor that is required to potentiate the protein to a toxic form. A form of potentiation has been observed previously in *Photorhabdus* and is best understood in the Tc toxin complexes (Waterfield *et al.*, 2005).

(A)

MSEIEAITLKAPISVIAKVT**N**KSVYKLKLIQDSIRLDQGEWTTLPPQVI**N**
RSNDNTPGKAIWRSDSNSILSGVAGRCTYVFVDSKGEIYSIYITWSNPLI
GSSSYSISTDYEGDDLHLSYTADVGNNPVVEYTIVT

(B)

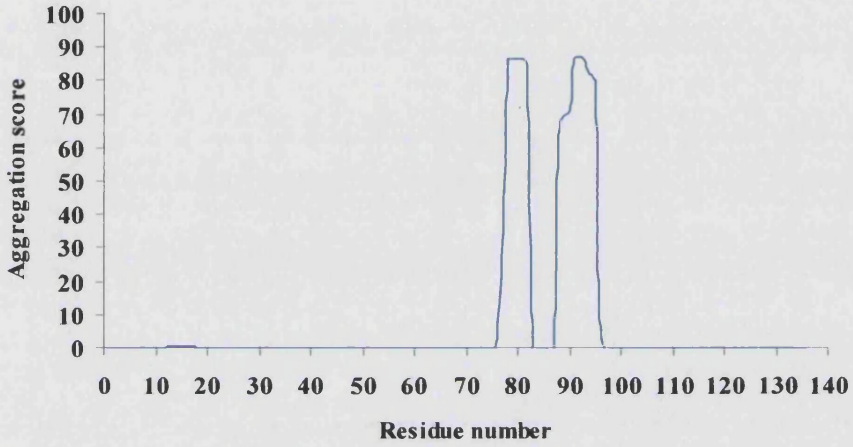


Fig. 2.24 Aggregation and glycosylation prediction for Mns. (A) Complete amino acid sequence of Mns with potential asparagine glycosylation sites indicated (bold). Amino acid residues implicated in β-aggregation have also been shown (blue) with possible gatekeeper residues (red) based on a TANGO plot of Mns **(B)**. No helix-aggregation was predicted by the server (See Chapter 3 Discussion).

The hypothesis that Mns is an insecticidal toxin comes from a BLAST analysis (Altschul *et al.*, 1997) of its primary sequence. The closest match produced is to the 14kDa subunit of the *B. thuringiensis* binary insecticidal protein, Cry34, with a peptide sequence identity of 27%. *B. thuringiensis* strains produce a variety of proteinaceous crystals, some of which are toxic to insects and other invertebrates, and some for which no biological activity has been detected. The binary insecticidal proteins, Cry34 (14kDa) and Cry35 (44kDa), are so-called because both proteins are required for effective mortality when tested on the western corn rootworm *Diabrotica virgifera virgifera* LeConte (Schnepf *et al.*, 2005). These toxins have been expressed from engineered genes in transgenic corn plants to provide high levels of protection against *Diabrotica* (Moellenback *et al.*, 2001). The insects show histopathological symptoms in the midgut epithelium, characterised by cell blebbing and vacuolation. Schnepf *et al.*, (2005) predict that the binary toxins disrupt membranes and that Cry34 may perform a membrane anchoring or translocation function. The authors identified two other proteins with a Cry34-like domain, each of which also contains a putative phosphatase or endo/exonuclease domain. As a result, it has been further suggested that the Cry34-like domain may be used for delivery of the secondary module, which then exerts its toxic or cell-regulatory effect (Schnepf *et al.*, 2005).

The method used here to test for toxicity has been used elsewhere, and was successfully applied to identify toxicity by oral administration of the *Photorhabdus* toxin complexes (Waterfield *et al.*, 2001). *Photorhabdus* are able to kill a variety of insects, and have a genome massively redundant in toxic factors with activity towards insects and bacterial competitors (Duchaud *et al.*, 2003). The Mcf toxin induces apoptosis of the insect midgut, while toxins of the *Photorhabdus* virulence cassettes cause destruction of insect haemocytes by rearrangement of the actin cytoskeleton. Analysis of the newly completed genome sequence of *P. asymbiotica* is uncovering yet more novel virulence factors (ffrench-Constant *et al.*, 2007). Because of this redundancy, we were unable to discriminate between the pathogenicity of wild-type and Mns-knockout cells.

It is worth noting that in addition to the role of altering cell surface properties, many secreted biosurfactants have been shown to have toxic activities. Their effect is achieved through membrane disruption, leading to cellular lysis and metabolite leakage, or through the interference of membrane functions. Interestingly, cationic surfactants are considered to be the most toxic (van Hamme *et al.*, 2006).

2.4.8 Mns Expression in Insects and Nematodes

The high level of secretion of Mns, and its presence throughout the infection process are consistent with the protein being a toxicity factor. Western blots of *G. mellonella* infected with *P. luminescens* and with *P. asymbiotica* strain ATCC43949 showed Mns detectable from the first day after infection. Mns from Kingscliff *P. asymbiotica* was first detected on the second day. *P. asymbiotica* ATCC43949 grown in liquid cultures produced detectable levels of Mns in the lag phase, which again would be consistent with the establishment of infection in the insect host. Whether the presence of Mns at day 15 of infection reflects continuous expression and requirement of the protein, or results from the persistence of protein expressed early on in the infection process, is unclear.

Mns also appears to be expressed in nematodes. The recently identified nematode hosts of Kingscliff *P. asymbiotica* were crushed and prepared for Western blotting, and a faint signal was detected with anti-Mns antibodies. However, symbiosis assays performed with Mns-knockout and wild-type *P. luminescens* have indicated that there are no differences in the transmission efficiency of infective juveniles transmitting either strain (T. Ciche, personal communication), suggesting that the protein plays no vital role in the *Photorhabdus*-nematode symbiotic lifecycle.

2.4.9 Conclusions

Our findings suggest that Mns is a secreted adhesive factor of *Photorhabdus* that modifies the attachment of cells to surfaces. In a natural environment, *Photorhabdus* are not expected to live freely and form biofilms on abiotic surfaces. Instead, the modification of cell attachment likely relates to insect colonisation, perhaps allowing adhesion in specific tissues. The interaction with *C. elegans* likely serves a different purpose. Since *C. elegans* is not a symbiotic host of any bacteria of the genus *Photorhabdus*, the clumping phenotype we observed could be a defence mechanism to avoid predation. Recently, we have observed that *P. asymbiotica* supernatant is able to aggregate *Cooperia oncophora*, *Teladorsagia circumcincta* and *Haemonchous contortus*, in addition to *C. elagans*, but not *Heterorhabditis bacteriophora*. Similarly, *P. luminescens* TT01 caused aggregation of *H. contortus*, but none of the strains tested were able to affect *H. bacteriophora*, the host species of *Photorhabdus* (N.R. Waterfield, unpublished data). As outlined in Chapter 1, bacterial interactions of this sort have been observed previously, with *Y. pseudotuberculosis* and *M. nematophilum* (Darby *et al.*, 2002, Hodgkin *et al.*, 2000). In both roles, cell-cell and cell-surface adhesion appear to be involved, high levels of secretion are likely to be necessary, and a high degree of conservation within the genus is consistent.

CHAPTER 3

The Structural Properties of Mns

3.1 INTRODUCTION

To further investigate the role of Mns, we found it of interest to purify the protein for structural studies. Samples were prepared for the high resolution techniques of NMR and crystallography, but the circular dichroism, dynamic light scattering and differential scanning calorimetric properties were also investigated in an attempt to characterise the protein and understand its role in biofilm formation and nematode aggregation.

3.1.1 Nuclear Magnetic Resonance

Nuclear magnetic resonance (NMR) and X-ray crystallography are the major techniques currently available for the determination of three-dimensional structures of proteins at atomic level. Of these, NMR allows the determination of structures of proteins in the solution phase and may be used to study protein dynamics. NMR is therefore invaluable in the study of protein structure and function.

3.1.1.1 Principles of Nuclear Magnetic Resonance

In an NMR experiment, data are collected from the nuclei of atoms. NMR-active nuclei, such as ^{1}H, ^{13}C and ^{15}N, have uneven numbers of subatomic particles and as they spin around an axis they possess an overall 'nuclear spin' that produces a magnetic field. The spin can have two possible orientations when exposed to an external magnetic field: aligned with, or against, the applied field (Creighton, 1993a).

These two states do not have the same energy and there is a bias towards the lower energy state, i.e. the nuclei that align with the external magnetic field. When an electromagnetic radiation pulse is applied to the nuclei, at the frequency that matches the difference in the two energy states, resonant absorption occurs, so 'flipping' (resonating) the magnetic moment of the nuclei into the higher energy orientation where it opposes the external magnetic field. As the nuclei relax to the lower energy state, they emit electromagnetic energy that can be recorded by a detector (Branden & Tooze, 1999). It is this 'signal' that is measured in NMR experiments.

As different nuclei resonate at slightly different frequencies, the electromagnetic pulse is applied at many radio frequencies simultaneously. The frequency of resonance is a function of the nucleus and the shielding effect of local currents produced by electrons around the nucleus. Binding partners, bond angles and close-contact interactions associated with the nucleus all modify the distribution of electrons around the nucleus and therefore alter the local magnetic field (Wishart & Case, 2001). This shielding effect is also known as the chemical shift and is plotted to provide details of the magnetic fields that exist in the sample. If the electrons that surround a nucleus provide a strong shielding effect that opposes the applied magnetic field, the chemical shift is said to move upfield on a one-dimensional NMR spectrum (**Fig. 3.1 (A)**).

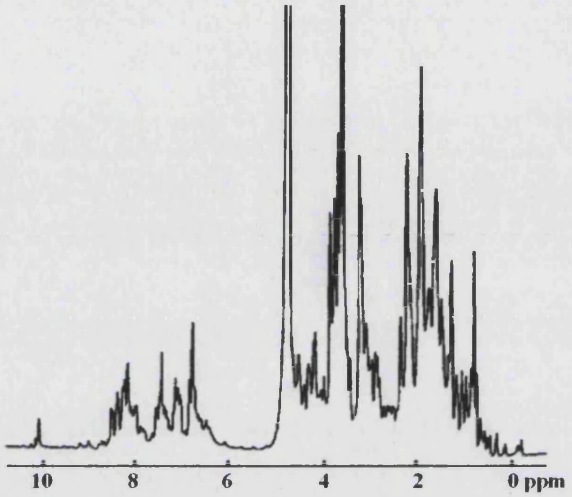
The chemical shifts of protons bonded to nitrogen are poorly shielded and appear 'downfield' on a ^{1}H-NMR spectrum. Conversely, protons associated with carbon atoms forming aromatic, aliphatic and methyl groups are increasingly shielded by electrons and appear upfield on the spectrum, representing higher resonance frequencies.

3.1.1.2 Sequential Assignment and Structure Determination

The quality of a one-dimensional spectrum, displaying all of the chemical shifts present in a sample, gives an indication of the suitability of the protein for NMR analysis. Unfolded proteins produce an NMR spectrum that is essentially the sum of all the amino acid residues, but specific folding of the protein results in highly dispersed signals reflecting the many different chemical environments in which the nuclei exist. Poor dispersion or broad spectral line widths suggest that the protein forms a higher molecular weight entity such as non-specific aggregates or large homo-oligomers. Similarly, dynamic events such as fluctuation between different conformations also cause poor sensitivity (Yee *et al.*, 2005).

One-dimensional ^{1}H spectra of proteins contain overlapping signals from many hydrogen atoms because, although most may produce a unique signal, some exist in environments that are chemically similar. Consequently it is not possible to assign each of the spin systems to particular hydrogen atoms in the protein.

(A)



(B)

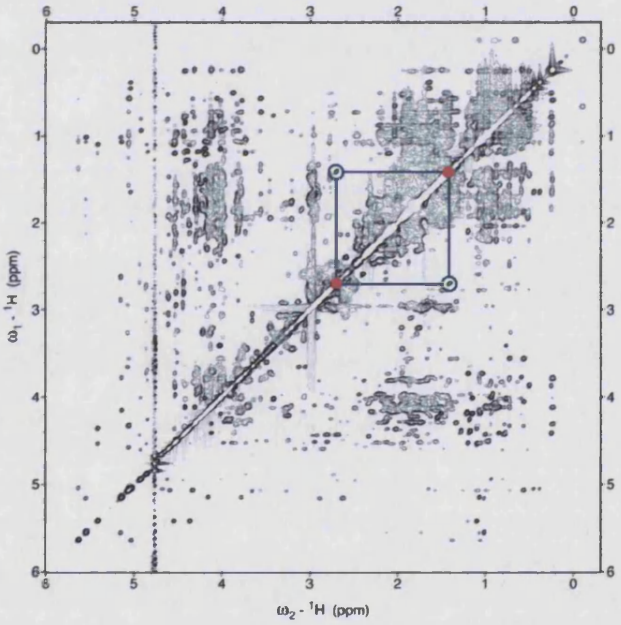


Fig 3.1 Typical one-dimensional and two-dimensional spectra. **(A)** One dimensional spectrum of murine prion protein (PrP) with chemical shift shown along the x axis in parts per million (ppm). (From Hornemann *et al.*, 1997) **(B)** Two-dimensional nuclear Overhauser effect spectroscopy (NOESY) spectrum of the chemotaxis inhibitory protein from *Staphylococcus aureus* (Chips) protein. The diagonal line (coordinates 6,6 to 0,0) represents a one-dimensional spectrum. Cross peaks lie off the diagonal line, and in NOE spectroscopy can be used to identify nuclei that are 5Å apart or less. In the highlighted example, the circled cross peaks show that nuclei in red are connected in this way. (Spectrum kindly provided by S. Bagby, University of Bath).

Two-dimensional NMR spectroscopy has been developed to overcome this problem. Here, a series of radio frequency pulses rather than a single pulse in used, with a delay period between them. Different NMR experiments use different delay periods, as well as different frequencies and pulse intensities, which in some cases allow the nuclei to freely precess for a determined time period. After the final pulse the frequencies of the nuclei are recorded. In correlation spectroscopy (COSY) experiments, frequencies from the nuclei are plotted along both axes, with the diagonal corresponding to a normal one-dimensional spectrum (Markley, 1989). Signals that lie off the diagonal are 'cross' peaks and represent coupled nuclei. By matching a cross peak to two corresponding peaks on the diagonal line, it is possible to determine which nuclei are coupled, and therefore which atoms are covalently connected to each through one or more covalent bonds **(Fig. 3.1 (B))**.

Because some amino acid residues have characteristic patterns of couplings and chemical shifts one can begin to assign well-resolved signals to particular residues from COSY data. It then remains to assign NMR signals to specific amino acids in the protein using the nuclear Overhauser effect (NOE), which gives peaks between pairs of hydrogen atoms separated through space by less than 5Å. NOE spectra therefore record interactions between hydrogen atoms in adjacent residues. With knowledge of the peptide sequence, and prior identification of some of the residues of the connected signals, it is possible to assign the remaining amino acids in sequence (Basus, 1989).

The NOE is also the most important source of conformational information in structure determination once all of the NMR signals have been assigned to the correct nuclei. In α-helices and β-sheets neighbouring protons of the peptide backbone are brought so close together that they can be observed by NOE spectroscopy and give characteristic patterns of cross peaks (Brandon & Tooze, 1999). Models of the three-dimensional structure of the protein are then derived using coupling constraints and chemical shifts that provide dihedral angle measurements. A molecular dynamics programme calculates many structures, each of which equally well obeys the structure constraints enforced by the NMR data. When there is sufficient quantity and quality of constraints derived from the NMR data, only one structure obeys all the restraints.

3.1.1.3 Size Restrictions

Successful NMR spectroscopy relies on high quality spectra, recorded with good sensitivity and resolution. It is more difficult to achieve these requirements as the size of the macromolecule increases and few studies report the solution structures of proteins larger than ~30kDa. A significant problem is signal overlap due to the complexity of the spectra, which may be overcome by segmental labelling (Wider, 2005). Alternatively, for resonance assignments of proteins larger than 40kDa, Takeuchi *et al.* (2007) have used selectively labelled amino acids.

The slow tumbling of large macromolecules leads to rapid relaxation of transverse magnetisation as a result of enhanced spin-spin interactions, causing line broadening and low sensitivity (Wider, 2005). The development of cryogenic probes has improved the sensitivity of NMR measurements, as has the use of transverse relaxation optimised spectroscopy (TROSY)-based NMR techniques. Split signals from heteronuclei have different relaxation properties and TROSY retains only the slower relaxing component of the split signals. This produces more simple spectra with narrow lines and higher sensitivity (Foster *et al.*, 2007). Since the majority of biologically interesting macromolecular complexes are significantly larger than 25kDa, these advances have extended the application range of NMR and allowed analysis of protein assemblies. In addition to a general requirement for proteins to be small, if it is to be analysed by NMR a protein must be soluble, stable and should not aggregate at the high concentrations necessary for data collection.

3.1.2 X-Ray Crystallography

X-ray crystallography is built upon the principle that the regular arrangement of atoms in a crystal produces an interference pattern on a photographic plate when light of the right wavelength passes through. Examination of the interference pattern allows the arrangement of the atoms in the crystal to be determined.

The first protein structure to be solved by X-ray crystallography was that of myoglobin from the sperm whale (Kendrew *et al.*, 1958). Now, close to 5,000 crystal structures are deposited in the Protein Data Bank each year. X-ray crystallography has also found particular use in the design of drugs, because the structure of a receptor or ligand can be analysed and suitable drugs can be developed to bind the target without the extensive random screening otherwise required. The method is also being applied in high-throughput efforts to determine the structure of at least one member in every large protein family (Doerr, 2006).

3.1.2.1 Crystallisation

For X-ray analysis a single, three-dimensional crystal of at least 0.1mm in all dimensions is required (Wiencek, 1999). High internal order is also necessary to ensure the crystal will diffract X-rays to a suitable resolution: diffracting to 2Å or less allows individual atoms to be distinguished within a large protein structure. It is still not understood why some proteins crystallise with ease while others do not form crystals of sufficient quality for X-ray analysis, and this remains the major barrier to structure determination. The problem arises (a) because proteins are very sensitive to external conditions, and (b) because a very large number of experimental conditions may need to be tested before a suitable condition for crystallisation is found (Chayen, 1998). Furthermore, there is a requirement for extremely pure material.

Generally it is considered that any protein that can be studied by NMR may eventually be coaxed to yield a crystal for X-ray analysis (Doerr, 2006). However, there is no correlation between the quality of the HSQC spectrum of a protein and the probability of obtaining suitable crystals for crystallography (Yee *et al.*, 2005). Two approaches may be taken to overcome the difficulties of obtaining crystals; (a) a study of the fundamental factors that influence crystal growth such as solubility and nucleation kinetics, and (b) an empirical exploration of many different solution conditions. The latter approach is more commonly used, and benefits from sparse sampling matrices that efficiently sample a range of precipitants and pH conditions for their ability to crystallise the protein (Weincek, 1999). High-throughput screens, which may sample several hundred conditions with a minimal amount of purified protein, greatly improve the scope of the experiment. For instance, the high-throughput crystallisation laboratory at the Hauptman Woodward Institute screens

over 1,500 conditions, and is currently compiling a database of successful conditions that will be used to predict suitable conditions for crystallising novel proteins (Luft *et al.*, 2001)

3.1.2.2 Crystal Screening

Crystallisation requires bringing a protein into a supersaturated liquid state. The degree of supersaturation determines the rate of nucleation and crystal growth, and factors that affect solubility also affect these processes. These factors include temperature, pressure and components of the solution such as electrolyte additives and pH, and can be used to control the rate at which the crystal grows (Wiencek, 1999).

The most widely used technique to screen the conditions that may allow crystal growth is vapour diffusion, using the 'hanging drop' method. A drop of solution containing the protein is mixed with an equal volume of a precipitant and placed on a glass cover slip. The cover slip is inverted over a reservoir containing the precipitant and because the solution in the reservoir contains no protein, it is at twice the precipitant concentration of the drop. Initially the drop is thus undersaturated, but over time water in the drop evaporates and equilibrates with the water in the reservoir. This changes the supersaturation state of the protein solution, with the protein becoming more concentrated and less soluble as the precipitant concentration increases.

In 'batch' crystallisation techniques supersaturation is reached immediately upon mixing the protein with the crystallising agents. This is the most common technique for automation and allows many different conditions to be screened. To reduce consumption of the sample, drops may be 1-2µl in volume and overlaid with a mineral oil to minimise evaporation (Chayen, 1998). D'Arcy *et al.* (1996) have found that with some mixtures of oil it is possible to shrink the drop over time, so mimicking the concentration effects valued in vapour diffusion.

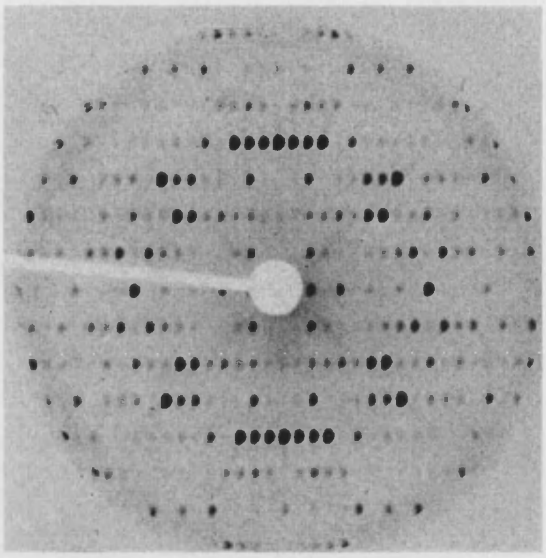
Finally, nucleation is the cooperative step-by-step addition of monomers, and requires extremely high levels of supersaturation. This phase, and that of crystal growth, are poorly understood, but are known to have different requirements. An approach to separating these stages is crystal seeding, which involves transferring a small crystal into conditions of lower supersaturation in which crystals will grow but no further nuclei will form (Chayen, 1998).

3.1.2.3 X-ray Diffraction and Data Collection

Crystals are composed of regular arrays of atoms that occur in repeated units within a three-dimensional lattice. Commonly, a single protein molecule makes up an asymmetric unit, which is the minimum structure necessary to reproduce the whole crystal. Two asymmetric units comprise a unit cell, which has been described as a single brick that repeats in three directions to build the crystal (Souza *et al.*, 2000). To prepare a crystal for crystallography, it is usually preserved in a cryoprotectant and frozen to liquid nitrogen temperatures. The crystal is then exposed to a narrow beam of X-ray radiation (Holmes, 1999).

X-ray beams at synchrotron sources are a product of circular electron accelerators and storage rings. An X-ray monochromator selects radiation at a single wavelength, and the crystal is placed in the path of the intense, tightly focused beam. The electric field of the X-radiation causes the electrons of atoms in the crystal to oscillate and emit scattered radiation. The scattering of X-rays, recorded at an electronic detector, produces a pattern of spots known as Bragg reflections (**Fig. 3.2**). Every atom in the crystal contributes to the scattering that produces these spots, and each spot receives some contribution from each atom (Souza, 2000). During data collection, the crystal is rotated relative to the angle of the incident beam, and a diffraction pattern is collected at 1° increments. When a complete set of diffraction patterns is collected, the data are indexed to identify the dimensions of the unit cell, and integrated using a computer programme such as the HKL package (Otwinowski, 1993). The result is two measurements for each spot: its position (recorded using the Miller indices with coordinates *h*, *k* and *l*) and its intensity. Lastly, scaling is necessary to remove differences in spot intensities incurred over time by radiation-induced damage and beam variations experienced over the period of data collection.

(A)



(B)

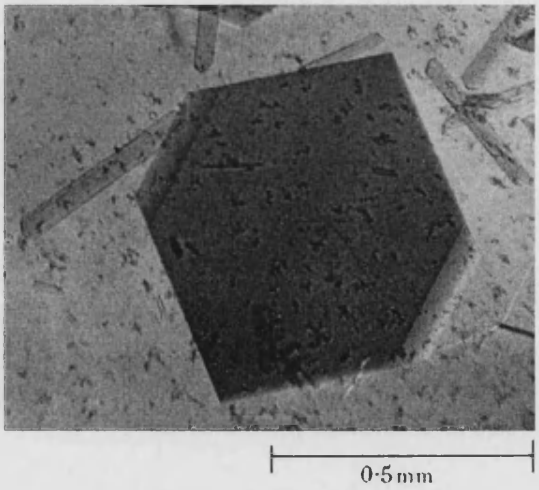


Fig. 3.2 Example diffraction pattern. **(A)** Precession photograph recorded on X-ray film. Each dark spot is a reflection produced when X-rays are scattered by electrons in the crystal. The pattern shown is that of a myoglobin crystal (Kendrew *et al.*, 1960). Analysis of the locations of the reflections allows the determination of the location of electron clouds within the crystal, and ultimately the structure of the protein. **(B)** Crystal used for collection of the diffraction data.

The position of diffraction spots is determined by the size and shape of the unit cell, and the symmetry within a crystal. Each reflection recorded during data collection is the product of a diffracted beam with three components, a wavelength, amplitude and phase. The wavelength is fixed in the experiment, and the amplitude can be measured by the intensity of the Bragg reflections. Unfortunately, as the diffracted ray hits the detector, phase information is lost but is needed to obtain an electron density map. This is known as the phase problem, and three approaches have been taken to solve it, including molecular replacement and isomorphous replacement. The development of synchrotron sources of radiation, with X-ray beams that can be tuned to different wavelengths, allows an alternative method, anomalous dispersion, to be used for phase determination.

Anomalous dispersion uses the property of anomalous scattering by atoms. This is a change in scattering intensity and phase produced when X-rays are used at a wavelength close to the absorption edge of an atom, and is caused by the inner electrons of an atom. As with isomorphous replacement, the crystal is usually prepared with electron dense atoms, because they make a greater contribution to the scattering pattern at the wavelengths required for data collection (Hendrickson *et al.*, 1990). By changing the wavelength of the X-ray beam the scattering strength can be varied. Collection of data at different wavelengths then allows the phases of the anomalous diffracting atoms to be obtained.

The introduction of selenium into crystals is now the method of choice for crystallographers needing heavy atoms in a crystal, and may be achieved by growing the protein on minimal media supplemented with selenomethionine (Hendrickson *et al.*, 1990).

3.1.2.4 Structure Determination

Once the phases are known, electron density maps can be obtained by Fourier synthesis. Electron densities are mapped in three-dimensions, and show differing density distributions through the use of contour lines. Electron dense areas correspond to atomic positions but due to limitations in resolution will not normally appear as isolated islands. The fidelity of the map depends on the resolution of the Fourier synthesis, and derives from the number of scattered intensities used (Souza *et al.*, 2000, Creighton, 1993a). Continuous stretches of electron density should be visible, and correspond to the polypeptide backbone. A map at 6Å resolution will reveal only the course of the chain, but atom groups may lie between 2.8Å and 4.0Å apart, and individual atoms are between 1.0Å and 1.5Å in size. To distinguish atoms, the best possible resolution is required, and will depend ultimately upon the quality of the crystal.

Maximum diffraction resolution also improves map interpretability, facilitating automated construction of an atomic model (González, 2003). Electron density maps can be manually and interactively interpreted with molecular graphics programmes, such as "O" (Jones *et al.*, 1991). The polypeptide chain is traced from either terminus or from a convenient starting point such as a disulphide bond, and the known amino acid sequence of the protein is built into the density map, ensuring well-known features of the backbone and sidechains are well preserved (Souza *et al.*, 2000).

Finally, errors in the model are minimised through 'refinement'. This involves adjustment of atomic coordinates of the model in an attempt to correlate the model with the experimentally derived diffraction data. Refinement is an iterative process, in which phase information is improved and fit against the observed diffraction data to improve the electron density map.

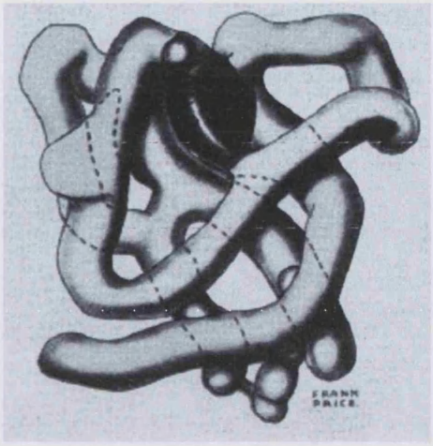
To reduce the number of parameters to be refined, certain atom positions can be fixed relative to each other if the bond lengths and angles that separate them have been fully characterised in other molecules. Errors can be corrected until correlation between the diffraction data and the atomic model are maximised. To correct large errors it may be necessary to manually rebuild the model with a molecular graphics programme. Simulated annealing may also be used, in which the protein is 'heated'

to up to 4000K and allowed to gradually cool. This simulated annealing allows large conformational changes in the protein and permits the automatic correction of large errors in the model (Souza *et al.*, 2000). During annealing, the model is held within the confines of the electron density map. As refinement continues, the electron density map is recalculated, based on better phases that result from the new positions the atoms adopt. The quality of the model may be assessed during refinement, using the R-factor and R-free parameters.

Recent advances in data collection have come from improvements to X-ray beam size and brilliance, and the development of charge-coupled detectors of sufficient size and resolution to replace X-ray film. Maintaining crystals at cryogenic temperatures allows collection of a complete data set from a single crystal, so removing the chance for crystal-to-crystal variations and minimising the opportunity for crystal damage during data collection. Statistical techniques for data modelling have also advanced.

Together these improvements allow higher quality and more accurate structures to be produced, as demonstrated in models of myoglobin (**Fig. 3.3**). Furthermore, these improvements enable the use of small (20-50μm) and weakly diffracting crystals, which for many proteins may be the only crystals obtainable, and have pushed the boundaries of high-resolution data (Beauchamp & Isaacs, 1999). Currently, the resolution of almost 8% of the structures submitted to the protein data bank are between 1-1.5Å, including some proteins to greater than 0.8Å, such as the *Bacillus* serine protease subtilisin (Kuhn *et al.*, 1998). The record resolution stands at 0.54Å for crambin, a protein from Abyssinian cabbage that has just 46 residues and no known function (Jelsch, 2000).

(A)



(B)

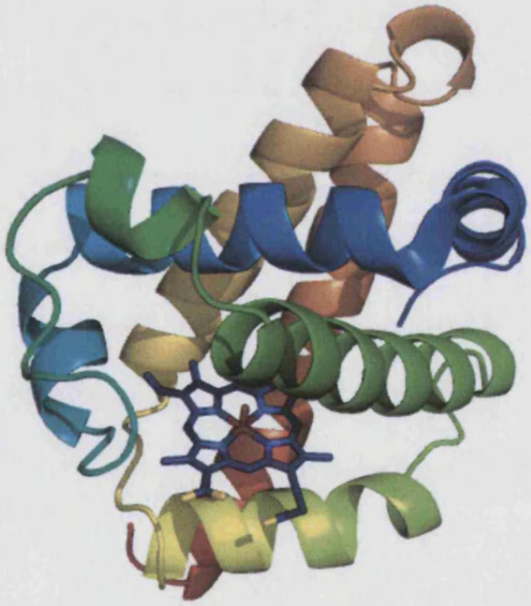


Fig. 3.3 Structure of sperm whale myoglobin. **(A)** Drawing of the structure of sperm whale myoglobin, as deduced from a 6Å Fourier synthesis, published by Kendrew *et al.* in 1960. **(B)** Improvements to crystallographic techniques and instrumentation, as well as to computational methods for modelling and model refinements, have improved the structural model of myoglobin and allowed its structure to be solved at 1.2Å resolution (Vojtěchovský *et al.*, 1999).

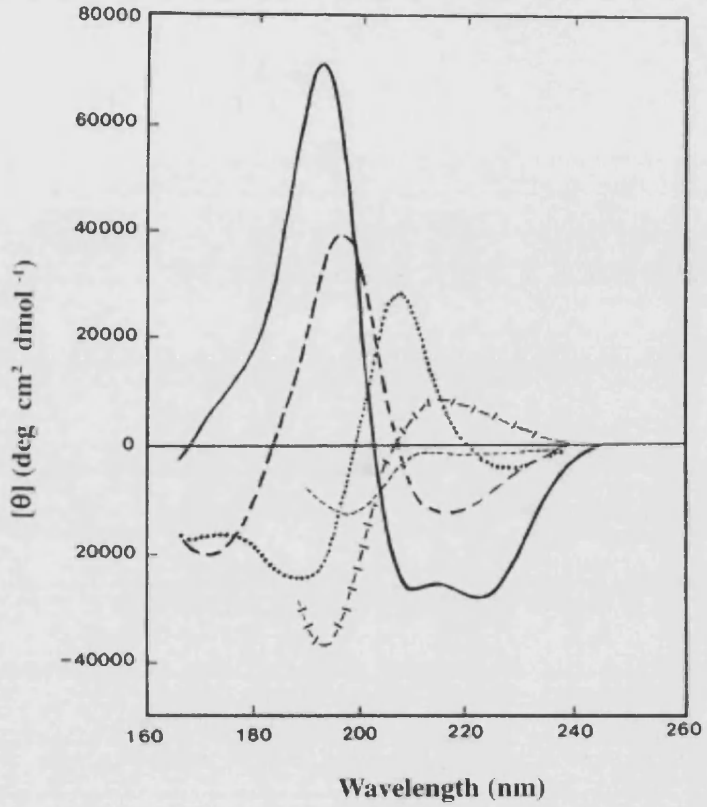
3.1.3 Circular Dichroism Spectroscopy

Circular dichroism (CD) spectroscopy can be used to analyse the different secondary structural elements exhibited by a protein. It is particularly useful for studying proteins in conditions close to those in which they usually 'operate', or for analysing the rate at which structural changes occur.

The technique is based on the differential absorption of plane polarised light, which is made up of two circularly polarised components of equal magnitude; right-handed light that rotates clockwise, and left-handed light that rotates anti-clockwise. A sample placed in a spectropolarimeter that is achiral will not produce a CD signal. However, samples with chirality are optically active and differentially absorb the two components of the light. This results in elliptical polarisation that can be detected as a spectrum when measured as a function of wavelength, with a positive signal when left-handed light is absorbed more than right-handed light, and a negative signal when right-handed light is preferentially absorbed (Reviewed by Kelly & Price, 2000).

The secondary structural components of proteins include α-helices and β-sheets, which are formed early in the protein folding process by local inter-residue hydrogen bonds. These components impart a distinct CD to the molecule. The spectral data of a protein can therefore be used to make predictions about the presence of these components. In the far-UV range, approximately between 260 and 180nm, α-helix, parallel and antiparallel β-sheet and turn structures give rise to characteristic spectral features and can all be identified (**Fig. 3.4**). Toward the near-UV, non-protein co-factors such as flavins and haem groups absorb radiation and can induce a signal. Aromatic amino acids and disulfide bonds also act as chromophores in this region, and the signals they produce may indicate whether the protein has folded into a well-defined tertiary structure. CD is highly sensitive to changes in conformation, and near-UV spectra are of particular use in monitoring these changes in response to the addition of a ligand or a change in buffer conditions. CD spectral data have been collected in the present study to gain an initial insight into the structure of Mns.

(A)



(B)

Secondary Structure	Positive Band	Negative Band
α helix	190nm	208 and 222nm
β strand	198nm	215nm
β turn	210nm	190nm
Random coil	212nm	195nm

Fig. 3.4 Typical Far-UV CD spectra. Spectra obtained from proteins with different secondary structures **(A)**. The characteristics of the curves are used to approximate the fraction of these components in test samples. Lines represent α-helix (solid); anti-parallel β-sheet (long-dashed line), type 1 β-turn (dotted line); extended 3_{10}-helix (cross-dashed line) and irregular structure (short-dashed line). The positive and negative maxima that define these curves are tabulated in **(B)**. (From review by Kelly *et al.*, 2005, van Holde *et al.*, 1998)

3.1.4 Thermal Analysis

Further to identifying the secondary structural elements of a protein, techniques are available to investigate the thermal stability of a sample, and the heat energy involved in physical processes such as denaturation and the establishment of interactions with other proteins or ligands. By heating a protein sample alongside a reference sample, it is possible to measure the heat energy taken up by the protein and begin to investigate the nature of the interactions.

Differential scanning calorimetry (DSC) uses a dilute solution of a protein of known concentration in a calorimeter that can be heated up to 130°C. The test protein and the control buffer are individually heated, but maintained at the same temperature throughout the experiment. The reference buffer provides an instrumental baseline for the experiment, and has a well-defined heat capacity (C_p) over the range of temperatures to which it will be exposed. As interactions are broken or formed with an increase in temperature, more or less heat must be supplied to the protein sample to maintain it at the same temperature as the reference (Cooper, 2001). This amount of heat gives a direct measure of the amount of energy absorbed or released during the transition (Cooper, 1999).

3.1.4.1 Protein Denaturation

Thermal unfolding of a protein in solution is usually an endothermic event; bonds are broken and the protein takes up energy during the transition. A typical DSC scan of this event from a physiological temperature towards boiling point follows a curve that peaks at the transition midpoint (T_m) (**Fig. 3.5 (A)**). At this point 50% of the protein molecules are folded and 50% are unfolded. If the transition is in an ideal reversible 2-state equilibrium this is the temperature at which any one molecule will spend 50% of its time in the folded state, and 50% unfolded (Cooper, 2001).

The compact globular state of a native protein is maintained by noncovalent interactions. These include ionic bonds, hydrogen bonds and the hydrophobic effect, which results from the weak attractive forces that bring non-polar residues together, and the stronger forces of repulsion that drive them away from the water medium. Additionally, some protein folds are held by strong disulphide bridges that form between cysteine residues (Whitford, 2005a).

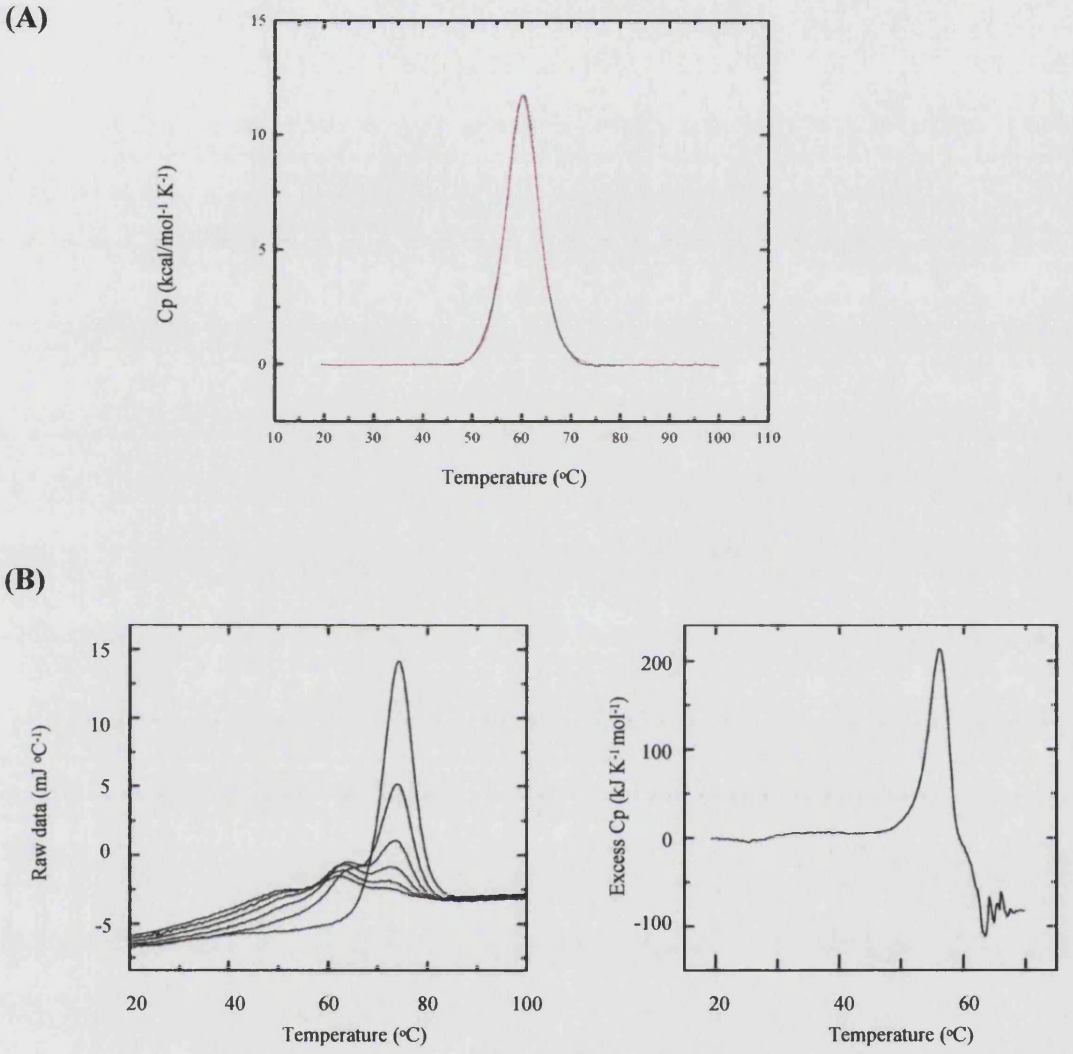


Fig. 3.5 Typical normalised differential scanning calorimetry curves. The denaturation occurs above approximately 50°C, and reaches the transition midpoint (T_m) at 62°C **(A)**. The red line shows the initial scan, and the black line shows the rescan after cooling. The protein follows a simple 2-state unfolding transition, and the protein fully refolded without aggregation. **(B)** Misfolded states are indicated by a shoulder before the major peak (left), and aggregation is characterised by a distortion of the curve with a exotherminc (negative) drop in energy (right). (Adapted from Cooper, 2001)

The physical properties of a folded protein usually show little change under physiological conditions but as it is exposed to increases in temperature, or concentration of a denaturant, a point is reached when the protein rapidly changes conformation towards the unfolded state (Creighton, 1993b). This is because the process is cooperative, and the denaturant acts principally on the denatured state; as long-range interactions are broken, the protein is weakened by an increase in entropy, and associated decrease in free energy, which destabilises the structure and exposes more residues to the denaturant (Shortle, 1996). Structureless coil regions are the first to unfold, and are followed by α and β structural elements that are stabilised by hydrogen bonds and have higher denaturation temperatures (Muthusamy *et al.*, 2000). Interestingly, regions that show high thermal stability when a protein is denatured have in some cases been identified as core residues that act as nucleation sites to drive and stabilise the folding process (Muthusamy *et al.*, 2000, Fan & Nan, 2007).

The 'unfolded' state of a protein is conventionally viewed as a simple structure of random coils in which no residue side chains can interact. However, differences in the energy states of unfolded peptides have been observed between wild-type and mutant forms, indicating that some side-chain interactions are present (Green *et al.*, 1992). Additionally, structural data comparing unfolded proteins with random coils have provided evidence that the unfolded state is composed of an ensemble of conformations with large amounts of residual structure (reviewed by Shortle, 1996). These conformations are known collectively as the denatured state, and are poorly characterised both because most proteins are insoluble when denatured, and the structures they form are highly flexible and dynamic, making direct structural analysis nearly impossible (Religa *et al.*, 2005).

3.1.4.2 Energy of Denaturation

The noncovalent interactions that preserve the secondary and tertiary structure of proteins in the native state contribute to its enthalpic energy, which is marginally greater than the entropic forces that try to destabilise the protein, ensuring a favourable net energy balance. As a protein is denatured, the noncovalent bonds are broken, and a completely unstructured polypeptide has no noncovalent interactions; only the covalent bonds that hold the residues together remain. This means the residues are not constrained to a highly ordered structure but have rotational freedom,

and the polypeptide can access many different conformational states (Doig & Sternberg, 1995). Denaturation toward this state is therefore associated with an increase in entropic energy, which relates to the natural tendency of thermal motion to disrupt order (Cooper, 1999). As temperature is increased, so the entropic forces in the protein increase and become greater than the change in enthalpy. At very high temperatures the entropy of the unfolded state dominates and favours unfolding (Whitford, 2005b).

A DSC thermogram records the energy of this unfolding as a change in heat capacity. This is defined as the change in enthalpy as a function of temperature, and has a positive value for protein unfolding because of an increasing exposure of non-polar amino acids to the medium (Whitford, 2005b, Sturtevant, 1977). Similarly, the changes that occur during protein-protein and protein-ligand interactions are associated with changes in heat capacity and can be measured by DSC. The enthalpy change of the reaction can be estimated from the thermogram by two independent means. Firstly, the calorimetric enthalpy (ΔH_{cal}) is given by the area under the transitional peak and is an absolute measure of the heat energy taken up by the transition. Secondly, the van't Hoff enthalpy (ΔH_{VH}) is a measure of the cooperativity of the transition, and depends on the shape of the curve. A large ΔH_{VH} will be recorded if the peak is sharp, indicative of a high degree of transitional cooperativity (Cooper, 2001).

3.1.4.3 Protein Refolding

After cooling, rescanning the sample is expected to produce a transition curve that resembles the initial scan, indicating the protein has been able to refold to its native state. Differences in the curve are obtained if a protein is unable to refold in the absence of a molecular chaperone, or if exposure of the peptide to high temperatures causes chemical changes that lead to improper folding. Lastly, aggregation of the unfolded polypeptide is common and may be irreversible, and is characterised on a DSC thermogram by an exothermic (negative) peak (**Fig. 3.5 (B)**) (Cooper, 2001). In the present study, recombinant Mns was analysed by DSC to study its thermal stability and its ability to refold to the native state after thermal denaturation.

3.1.5 Dynamic Light Scattering

Another technique available for the study of proteins in solution is dynamic light scattering (DLS). The technique is based on the interactions that take place between light and matter. Biological molecules scatter light, and do so at an intensity that is dependent on their concentration and hydrodynamic diameter. Measuring the pattern of scattered light therefore allows estimates to be made about these properties, allowing predictions to be made on the dynamics of the sample and, when exposed to an electric field, its charge (Ford, 1985).

DLS, also known as quasi-elastic light scattering or photon correlation spectrometry, uses the Brownian motion of particles in solution to estimate their size. Brownian motion is the random movement of particles due to bombardment by the solvent molecules that surround them (Pusey & Tough, 1985). This motion causes fluctuations in the intensity of a scattering signal when light is applied to the sample, and because large particles move more slowly than small particles, the patterns of fluctuation they produce will be different.

The light scattering signal fluctuations recorded by DLS instruments are fed into a correlator that compares the scattering intensity at successive time intervals to measure the rate at which the intensity varies. If a light intensity signal recorded at one time point is identical to that at a second time point the two are said to have perfect correlation and a unity of 1.00. As successive readings are taken from a random source, the correlation will decrease until no correlation exists. This is indicated by a unity of 0.00. If a sample contains large particles the intensity signals will change slowly and the correlation will persist for a long time. Small particles move rapidly and the signal received from the light they scatter will lose correlation more quickly (Malvern literature, 2006). These principles explain the shape and interpretation of DLS correlograms; the time taken for the correlation to decay gives an indication of particle size, and the angle of the decay gives an indication of the polydispersity of the sample.

The particle sizes given by DLS are calculated from the Brownian motion velocities in terms of hydrodynamic diameters. This is because the diameter is estimated from the diameter of a sphere that shows the same velocity when diffusing through a fluid. Any changes in the hydrodynamic properties of proteins in a sample, e.g. changes in conformation, will affect their diffusion speed and will be reflected in the hydrodynamic diameter estimate. DLS is very sensitive to these changes, and to alterations to the surface structure of a protein. The technique is non-invasive, requires only small amounts of material and has been used to study many biological occurrences including processes of fibrin polymerisation, the velocity of motile cells and recently the formation of a 2-oxo acid dehydrogenase multienzyme complex from its individual subunits (Bloomfield, 1985, M.G. Posner, personal communication). DLS has been used in the current study to analyse the conditions in which Mns forms aggregates, with the aim of reducing the occurrence of this phenomenon during NMR spectroscopy.

3.2 MATERIALS AND METHODS

3.2.1 Purification of Mns

Mns was purified using ion exchange chromatography and gel filtration chromatography as described in Chapter 2. For NMR, the protein was concentrated in 5mM K_3PO_4 pH 7.0 to 1mM in a sample size of 500µl.

3.2.2 Nuclear Magnetic Resonance

3.2.2.1 Unlabelled Mns for NMR

Initial NMR data were collected from unlabelled Mns. To remove the background signal associated with the hydrogen atoms of water, Mns suspended in HEPES pH 7.0 was exchanged into a deuterium buffer (5mM K_2HPO_4, adjusted to pH 7.0 with DCl/NaOD, 1mM NaN_3). A 5000MWCO centrifuge filter concentrator (Millipore) was washed with water then equilibrated with deuterium buffer. To a 200µl sample of purified Mns, 2ml of D_2O buffer was added and concentrated to 200µl by centrifugation at 800x g, 4°C. This was repeated twice more and the protein was collected. The filter was washed 3 times with 50µl D_2O buffer to collect the remaining protein.

3.2.2.1 Uniformly ^{15}N- and ^{13}C-labelled Mns for NMR

To produce Mns with ^{12}C and ^{14}N replaced with ^{13}C and with ^{15}N, the *mns*-pBAD30 vector was transformed into BL21 Gold (Novagen) cells, which allowed growth in minimal medium (1ml 1M $MgCl_2$, 0.1ml 1M $CaCl_2$, 0.1ml metal solution (550mg $CaCl_2.H_2O$, 140mg $MnSO_4.H_2O$, 40mg $CuSO_4.5H_2O$, 220mg $ZnSO_4.7H_2O$, 45mg $CoSO_4.7H_2O$, 26mg $Na_2MoO_4.2H_2O$, 40mg H_3BO_4, 26mg KI, 500mg EDTA, adjusted to pH 8.0, 375mg $FeSO_4.7H_2O$, made to 100ml and autoclaved), 4mg biotin, 4mg thiamine, 1g NH_4Cl, 2g glycerol, made to 1litre). For a ^{15}N-labelled sample, 1g of NH_4Cl per litre was used as the sole nitrogen source. For ^{13}C labelling, ^{13}C-glycerol (Spectra Stable Isotopes) was used as the sole carbon source (2g per litre). 2ml of an overnight culture was used to inoculate 200ml LB amp, which was grown to an OD_{600} 0.6 and pelleted. The pellet was used to inoculate 1l minimal medium, which was induced at OD_{600} 0.6 with L-arabinose (0.2%) and incubated at 37°C, 180rpm overnight. SDS-PAGE was used to confirm expression. Cells were pelleted

and Mns was purified with HiTrap QFF and MonoQ anion exchange chromatography as described. To prepare for NMR, pooled fractions containing Mns were concentrated to 1mM and exchanged into 5mM K_3PO_4 or 20mM HEPES pH7.0 using a 5,000MWCO filter concentrator. Buffer solutions were supplemented with 1mM NaN3, 1mM DTT, 1mM benzamidine and 10% D_20.

3.2.2.3 NMR Experimentation

NMR experiments were performed on a Varian Unity Inova 600MHz NMR spectrometer at the University of Bath. Data were collected on unlabelled samples using ^1H-^1H NOESY and ^1H-^1H TOCSY with mixing times of 100ms and 175ms (NOESY), and 50ms (TOCSY). The data were acquired at 25°C and then at 37°C to improve resolution. Data from Mns labelled with ^{15}N were collected using ^1H-^{15}N HSQC at 25°C, 37°C and ^{15}N-edited NOESY HSQC (100ms mixing time) 37°C, and 15N-edited TOCSY HSQC (50ms mixing time), 37°C.

3.2.2.4 Attempts to improve to NMR Data Quality and Removal of Metals

As Mns NMR samples tended to show significant reduction in data quality over a period of days, most likely due to protein aggregation, efforts were made to remove potential aggregation nuclei: preparations of sonicated cells containing Mns were centrifuged at 60,000x g for 5h at 16°C. The protein was also prepared to different degrees of purity and a crude sample was analysed in case the presence of other factors in the sample affected aggregation of the protein.

To remove metals from the sample, chromatography fractions containing Mns were treated with EDTA at a final concentration of 50mM and incubated at RT for 30min with gentle agitation. As before, potential aggregation nuclei were removed by centrifuging at 60,000x g for 5h at 16°C. Metals were removed from 5mM K_2HPO_4, pH 7.0, by treatment with 0.5g Chelex 100 resin (Bio-Rad)/100ml buffer. Metals were allowed to bind to the resin over 1h with gentle agitation. The resin was pelleted by 600x g, 10min centrifugation and the buffer was filtered with a 0.22μm syringe filter. Purified Mns was concentrated and buffer exchanged into the treated buffer using a 5000MWCO filter concentrator. The sample was made up to 500μl for NMR.

Lastly, a double-labelled sample of purified Mns was prepared for NMR with a Varian 800MHz spectrometer fitted with a cryoprobe at the MRC Biomedical NMR Centre, National Institute for Medical Research, Mill Hill, London. The protein was concentrated only to 0.5mM in 5mM K_2HPO_4, pH 7.0.

3.2.3 Crystal Trials

Purified Mns in 20mM HEPES pH 7.0 was concentrated with a 5000MWCO centrifuge filter concentrator (Millipore) at 800x g, in 10min runs at 4°C, and the optical density was measured at 280nm against a no-protein standard. The concentration of the sample was calculated by dividing the OD_{280} by the product of the extinction coefficient divided by molecular weight.

Purified protein, concentrated to 6mg/ml, was tested in a screen of 50 different formulations (Hampton Research Crystal Screen Kit) to identify initial crystallisation conditions. The trial was set up using the hanging drop method, with 1µl of protein added to 1µl of reservoir buffer on a siliconized glass circle cover slide (Hampton Research). Each was placed onto a grease-sealed well over 700µl reservoir solution. All crystal plates were stored at 18°C. The crystallisation condition that produced the largest crystals was 100mM Tris pH 8.5, 8% PEG 8000. Further screens were set up around this formulation, changing the pH of the buffer, the concentration of PEG precipitant and concentration of the protein.

3.2.3.1 Metal Additives

Metal salts were also added to crystal drops to improve crystallisation and examine the possibility that Mns is a metal binding protein. The metal chlorides and sulphate were added in the drop to final concentration of 10mM and 1mM (1µl protein:0.8µl reservoir buffer:0.2µl metal solution). No-metal additive and no-protein controls were also prepared. The metal salts used were $CdCl_2$, $CaCl_2$, $CuCl_2$, $MnCl_2$, $FeCl_2$ and $ZnSO_4$.

3.2.3.2 Micro- and Macro-seeding

Micro-seeding was performed with Mns crystals using a Seed Bead kit (Hampton Research). Small crystals from a drop of 100mM Tris pH 8.5, 8% PEG 8000 were removed using a mounted cryoloop and placed in a microcentrifuge tube containing 50µl stabilising solution (100mM Tris pH 8.5, 8% PEG 8000) and a Teflon Seed Bead. The tube containing the crystal and Seed Bead was vortexed for 90s and 450µl of the stabilising solution was added by pipette. This $1x10^{0}$ stock (dropping solution 1) was serially diluted by mixing 5µl with 45µl of stabilising buffer containing no protein, to give a $1x10^{-1}$ dropping solution 2. Each was further diluted to a final solution of $1x10^{-6}$ (dropping solution 7).

To set the drops, 1µl samples of each dropping solution were mixed with 1µl resevoir solution (100mM Tris pH 8.5, 6 or 8% PEG 8000 with or without 2mM $MnCl_2$ supplement). All drops were hung over a 700µl reservoir solution and were incubated at 18°C. Uncrushed crystals were seeded directly into fresh drops using a mounted cryoloop. Drops were prepared 24h and immediately prior to seed introduction.

3.2.3.3 High-Throughput Crystallisation Screening

Mns was prepared to a concentration of 11mg/ml for the Hauptman Woodward Institute (Buffalo, NY, USA, www.hwi.buffalo.edu) high-throughput crystallization screen (Luft *et al.*, 2003). The protein was purified with two steps of HiTrapQ FF ion exchange chromatography, and the buffer was exchanged during concentration with 20mM HEPES pH 7.0. Sodium azide was added to the protein to a final concentration of 0.02% and 850µl was sent to the HWI. The protein was kept at 4°C.

HWI prepared the crystal growth screen in a 1536 well microbatch-under-oil experiment, using 0.2µl of Mns and 0.2µl of cocktail solution in 1.85mm-diameter wells. Digital images of each well were taken at 5 time points, including one prior to addition of the protein, and manually viewed using MacroScope software.

3.2.4 Circular Dichroism and Differential Scanning Calorimetry

Purified Mns was prepared to a concentration of 1mg/ml in 5mM phosphate buffer, pH 6.0, and was analysed with CD spectroscopy by Sharon Kelly at the Department of Chemistry, University of Glasgow. The sample of 925µl was also used for DSC experiments, which were performed on a Microcal VP-DSC spectrometer at the BBSRC Microcalorimetry Service (Department of Chemistry, University of Glasgow).

3.2.5 Secondary Structure Prediction

The amino acid sequence of Mns was analysed for secondary structure elements using the SSpro program (SCRATCH server, University of California, Cheng *et al.*, 2005, Jufo (Vanderbilt University, Meiler & Baker, 2003) and PredictProtein (Columbia University, Rost *et al.*, 2004). To search for known protein domains, the Pfam programme (Sanger Institute, Finn *et al.*, 2006) was used.

3.2.6 Dynamic Light Scattering

Mns purified as above was concentrated to 1.0mg/ml and analysed by dynamic light scattering (Malvern Nano-S instrument) after being used for NMR data collection. 20µl was analysed in a 1cm pathlength cuvette. The sample was then filtered with a 0.022µm filter (Millipore) and reanalysed with DLS.

3.2.7 Analytical Gel Filtration

Mns was purifed by ion exchange chromatography, filtered with a 0.22µl syringe filter (Millipore) and concentrated to 5mg/ml. 500µl of protein was injected onto a Superdex 75 (GE Healthcare) gel filtration column. The protein was passed through the column at 1ml/min and 1ml fractions were collected. The elution profile was measured at 280nm. The column was calibrated with reference proteins (Sigma) of known molecular weight at the same flow rate and a calibration curve was drawn with Rm (Relative mobility) vs log of molecular weight. An estimated molecular weight of Mns was calculated using the calibration curve. SDS-PAGE was used to show which fractions contained Mns, and the corresponding elution peak gave an estimate of the size of recombinant Mns. The reference proteins used were aprotinin (6.5kDa), carbonic anhydrase (29kDa), bovine albumin (66kDa), alcohol dehydrogenase (150kDa) and blue dextran (2000MDa).

3.2.8 Congo Red Absorbance

To test a hypothesis that Mns may form amyloid-like fibrils, 50µl of purified Mns concentrated to 4mg/ml was incubated with 10µM Congo red (CR), following the method of Chapman *et al.*, (2002). Absorbance of the sample was measured between 400nm and 650nm.

As a positive control, prion protein PrP prepared to the same concentration was also incubated with the dye and absorbance was measured in the same range. Purified murine PrP was the kind gift of K.M. Uppington, University of Bath. A negative control was prepared with CR with 20mM HEPES, pH 7.5.

3.3 RESULTS

3.3.1 Nuclear Magnetic Resonance

3.3.1.1 Unlabelled Mns

Unlabelled Mns prepared in deuterium buffer was initially analysed by one-dimensional ^{1}H NMR spectroscopy. The 1D NMR spectrum showed the presence of several upfield shifted methyl group signals below 0 ppm, indicative of a folded protein (**Fig. 3.6**). This was followed by the acquisition of 2D ^{1}H-^{1}H NOESY and ^{1}H-^{1}H TOCSY spectra.

The TOCSY data show connections between protons connected by covalent linking, and the NOESY data show connections between protons that are separated in space by 5Å or less. The quality of both TOCSY and NOESY spectra were improved by increasing the temperature at which the data were collected from 25°C to 37°C. Following these initial 2D experiments, one-dimensional spectra were recorded to verify that the protein had not precipitated during exposure to 37°C over up to two-days data collection. The spectra were very similar to the pre-experimental spectra.

3.3.1.2 Uniformly ^{15}N- and ^{13}C-labelled Mns for NMR

^{15}N-labelled Mns was prepared in minimal medium containing $^{15}NH_4Cl$. HSQC spectra were recorded at 25°C followed by HSQC and a 3D ^{15}N-edited NOESY at 37°C. The ^{15}N HSQC spectra showed excellent peak dispersion, indicative of a high β-sheet content in the protein. However, examination of the HSQC spectra recorded after the 3D NOESY data collection showed a drastic loss of signal, indicating that the protein had precipitated, aggregated or degraded (**Fig. 3.7**). Repeated experiments were abandoned during data collection due to loss of signal intensity, and a ^{13}C ^{15}N double-labelled sample also suffered from severe loss in signal-to-noise after four days of NMR data acquisition at 37°C. When signal intensity was lost, there was no visible sign of precipitation. SDS-PAGE showed there was no degradation of the protein, and the presence of aggregation was confirmed by dynamic light scattering.

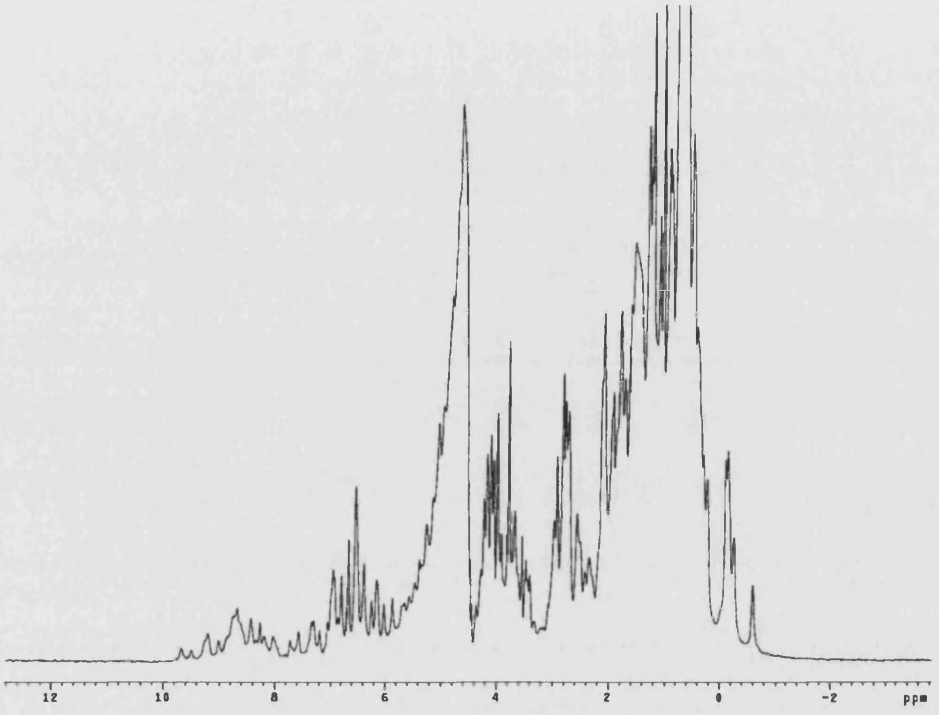


Fig. 3.6 One-dimensional NMR spectrum of Mns. Unlabelled Mns was prepared in phosphate/D_2O buffer. Peaks below 0 ppm indicate that the protein is in a folded conformation. Furthermore, the peaks are generally well defined and the line width is low, suggesting the protein is amenable to NMR spectroscopy.

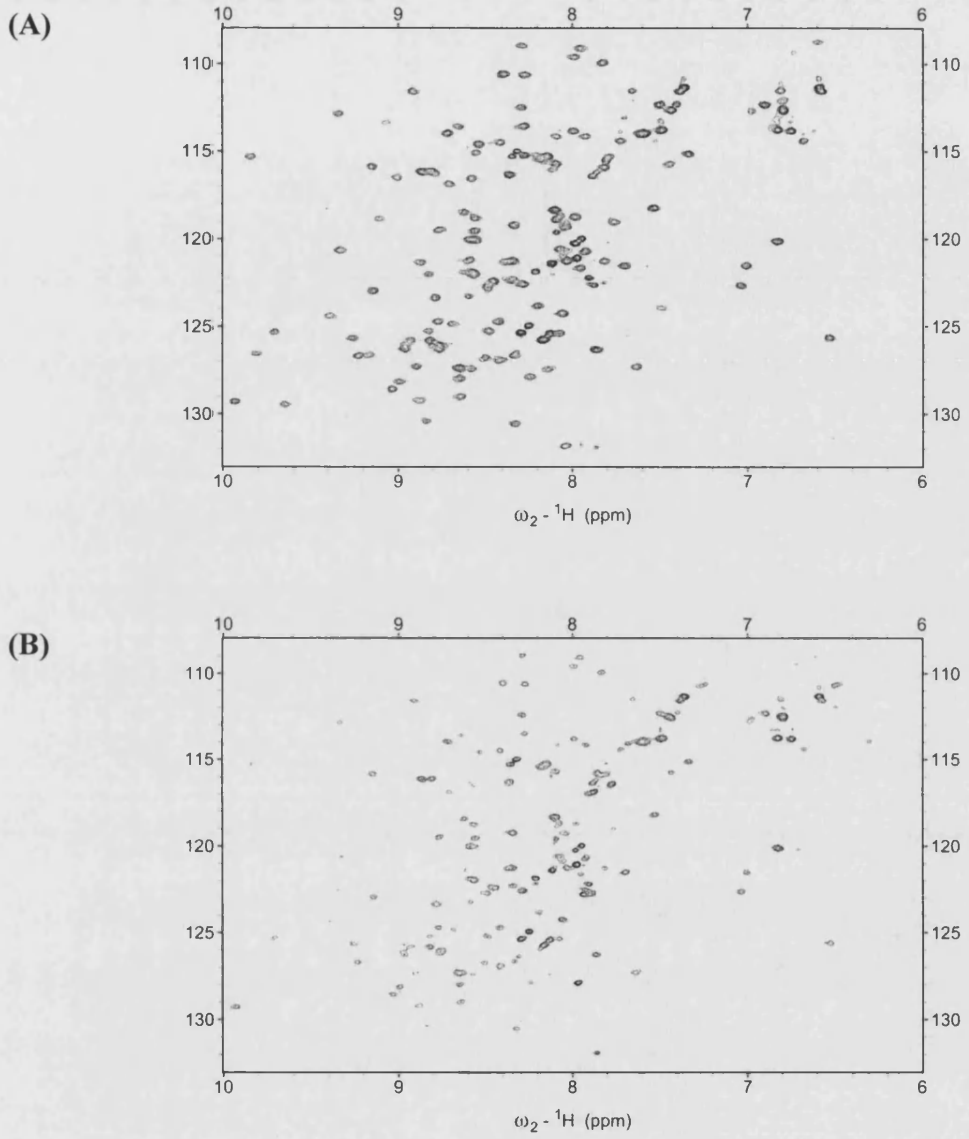


Fig. 3.7 Signal loss during NMR data collection. Two-dimensional ^{1}H-^{15}N HSQC spectra of ^{15}N-labelled Mns, collected in 30min at 37°C. Signal quality diminished over time from the initial spectrum **(A)** to a spectrum recorded three days later **(B)**. This signal loss was due to protein aggregation and prevented structure determination of Mns by nuclear magnetic resonance.

Dynamic light scattering analysis was used in an effort to improve data collection through reducing aggregation. Only by filtering a sample through a 0.022μm filter were we able to produce a monodisperse sample without aggregates (see below). Unfortunately the dilution required to achieve this state meant the sample could not be used for NMR. The majority of NMR data were collected at a concentration of 1mM. Samples at 0.75mM and 0.5mM were also used but failed to reduce aggregation, which occurred within two days of concentration. Furthermore, ultracentrifugation to remove potential aggregates and aggregation nuclei failed to improve the amenability of the protein to NMR.

Mns was exchanged into phosphate and HEPES buffers at pH 7.0 for NMR data collection. Thermal analysis showed we were also able to vary the temperature without risk of precipitating the protein (see later in this chapter). Gel filtration was tested as an alternative purification method, and some single residue mutants of Mns that did not show cell elongation and clustering when expressed in *E. coli* were also purified. None of the conditions tested improved the state of the protein and aggregation was still observed after the protein was concentrated for NMR, either immediately or after days of NMR data acquisition. To explore the possibility that Mns is a metal binding protein, and that metal binding may influence the propensity of the protein to aggregate, a sample was prepared in the presence of EDTA and suspended in a Chelex resin-treated buffer. The ^{1}H-^{15}N HSQC of metal-free protein contained about 40% fewer peaks than the untreated protein, but this was not reproducible.

Finally, data were collected using a cryoprobe Varian 800MHz NMR spectrometer at the MRC Biomedical MNR Centre, London. Despite the roughly 4:1 improvement in signal-to-noise ratio afforded by the cryoprobe and the sensitivity of the 800MHz spectrometer, extensive aggregation meant that no useful data could be collected.

The NMR data obtained for Mns were insufficient to provide any insight into the structure of the protein. Backbone assignments were started but the tendency of Mns to aggregate and failure to overcome this problem meant that NMR data of sufficient quality could not be acquired.

3.3.2 Crystallisation

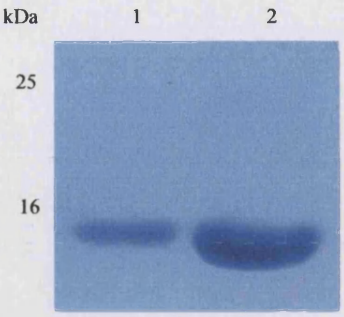
3.3.2.1 Crystallisation Screen

Prior to the crystallisation screen, Mns was purified by ion exchange chromatography, concentrated and passed through a gel filtration column as a final polishing step (**Fig. 3.8**). Approximately one third of the conditions tested resulted in precipitation. The most promising crystals came from condition 36 (8% PEG 8000, 0.1M Tris HCl pH 8.5) of the Hampton Research crystal screen kit, with some microprecipitation from conditions 27 (20% isopropanol, 0.1M Na, HEPES pH 7.5, 0.2M sodium citrate) and 42 (20% PEG 8000, 0.05M potassium phosphate). For optimisation, a screen was set up with 3%, 5%, 8%, 12% and 15% PEG 8000, using a pH range of 8.0 and 8.5 (0.1M Tris HCl) and pH 9.0 (0.1M Bicine). The largest crystals obtained came from macro-seeding a single crystal from an initial preparation in condition 27 into the same condition (20% isopropanol, 0.1M Na HEPES pH 7.5, 0.2M sodium citrate) containing 1µl of protein (5mg/ml) added immediately before the seed. Unfortunately, this showed the diffraction pattern of a salt crystal (data not shown).

3.3.2.2 High-throughput Crystallization Screen

A sample of Mns was combined with 1536 chemical cocktails in a high-throughput crystallisation screen at the Hauptman Woodward Institute. As part of the screen, each well was photographed prior to the addition of protein to the cocktail, and 1 day, 1, 2, 3, and 4 weeks after setup. The digital images were examined manually using MacroScope. In the high-throughput screen very small sample volumes are used and each well contained just 0.2µl of protein and 0.2µl of cocktail solution immersed in oil.

(A)



(B)

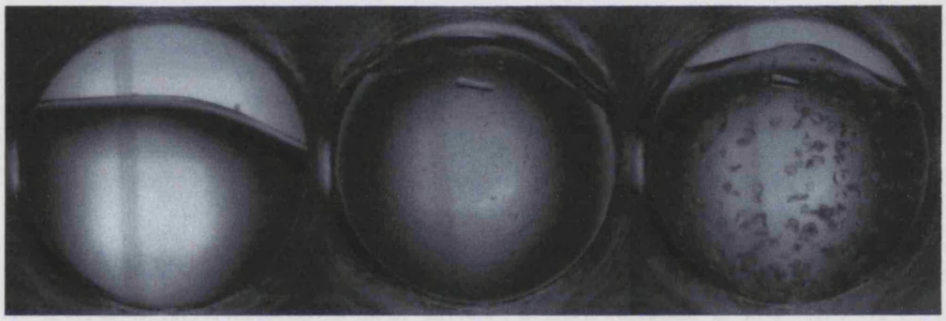


Fig. 3.8 Purified Mns and crystallisation screen. **(A)** SDS-polyacrylamide gel of Mns purified by ion exchange chromatography and gel filtration. The protein here was concentrated to 2mg/ml, with 2µl loaded in lane 1, 10µl in lane 2. The protein was further concentrated for crystallisation screens. **(B)** Compiled MacroScope images of Mns crystals grown in 0.79M zinc acetate, 0.1M sodium acetate, pH 5.0 (Hauptman Woodward Institute high-throughput crystallisation screen condition C0199). The image shows the drop before introduction of protein (left), at one day (middle) and after 1 week (right). The crystals developed no further and were identical after 4 weeks. The well diameter was 1.8mm and the largest crystals were approximately 100µm in size.

The largest crystals were obtained from condition C0199 (0.79M zinc acetate, 0.1M sodium acetate, pH 5.0) and were no larger than 100μm in size (**Fig. 3.8**). In-house reproduction of this cocktail mix failed to yield the quality of crystals observed in the screen. Condition C1405 (0.2M proline, 0.1M HEPES pH 8, 10%(w/v) PEG 3350) also produced very small crystals, and several others produced many thin needles (C1284 (0.1M Tris hydrochloride pH 9, 8 (w/v) PEG 8000), C1344 (2%(v/v) Dioxane, 0.1M Bicine pH 9, 105(w/v) PEG 20000)) or promising microcrystals and microprecipitation (C0201 (0.39M zinc acetate, 0.1M sodium acetate pH 5), C0202 (0.39M zinc acetate, 0.1M MES pH 6), C0265 (0.1M manganese chloride, 0.1M sodium citrate pH 4, 20%(w/v) PEG 20000), C960 (0.1M sodium molybdate dihydrate, 0.1M Tris pH 8, 60%(v/v) PEG 400)). As with condition C1099, these conditions were reproduced in-house in a vapour diffusion setup but failed to produce crystals.

3.3.3 Circular Dichroism of Mns

A 1.0mg/ml sample of purified Mns was analysed by CD spectroscopy at the University of Glasgow. Dilution of the sample in 5mM phosphate buffer, pH 6.0, to 0.383mg/ml was necessary to reduce the signal-to-noise ratio when measuring at the far-UV wavelengths (260-180nm), and data were collected from a 0.02cm pathlength cell. With these experimental conditions, the high tension voltage (HTV, the voltage applied to the system and used as a measure of absorbance) was kept below 550V, an acceptable level for the collection of reliable data.

The CD data of Mns was analysed using the CDSSTR variable selection method at the Dichroweb server (Whitmore & Wallace, 2004, Lobley *et al.*, 2002). Reference spectra set 3 (Birkbeck Crystallography Dept), covering wavelengths 240-185nm, gave the most consistent results when the Mns data were put into the reference set and the analysis was iterated. A normalised root mean standard deviation (NRMSD) value of 0.011 was achieved.

Far-UV spectra gave an approximate measure of the fraction and type of secondary structural elements the protein forms (**Fig. 3.9**). Mns is predicted to be largely helical, with 58% of residues involved in α-helices. These elements produce a strong positive maximum at 190nm, and negative maxima at 208nm and 222nm. CDSSTR is able to differentiate between 'regular' α-helices and 'distorted' helices, which do not have ideal phi and psi bond angles (Sreerama *et al.*, 1999). Similarly, 'regular' and 'distorted' β-strands have been identified (**Fig. 3.10**).

The maximum at 222nm is low in magnitude, and a small negative maximum exists at 215nm, both of which indicate that Mns has some β-strand component. CDSSTR analysis predicts only 8% of Mns residues form β-strands and 16% form turns. The remaining 18% of the protein would be disordered (**Fig. 3.10**).

CD data were also collected in the near-UV range (**Fig. 3.11**). Chromophores in this region, 350-250nm, include the aromatic amino acids and disulphide bonds, and can provide information on the tertiary structure of the protein. A spectrum was collected at 0.773mg/ml in a 0.5cm pathlength cuvette at a satisfactory HTV level (<320mV at the lowest wavelengths in the near-UV range). Less than 10% of the amino acid residues of Mns have aromatic side chains and the spectra collected were weak compared to the far-UV spectra. The data reveal little about the tertiary structure of the protein.

3.3.4 Secondary Structural Predictions

The SSpro program is the most accurate server available for the identification of secondary structural elements and predicted that 43% of Mns residues form strands and just 5% form helices (**Fig. 3.12**). Other servers also found relatively high levels of β-strand composition: 38% (Jufo) and 55% (PredictProtein). A search for known protein domains found none that were higher than the gathering threshold.

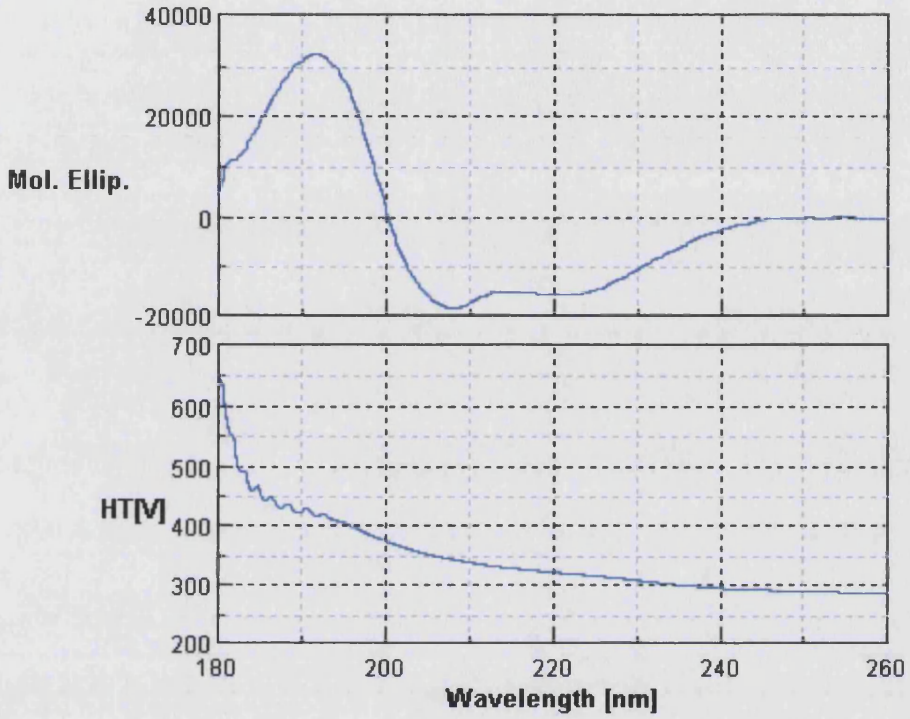
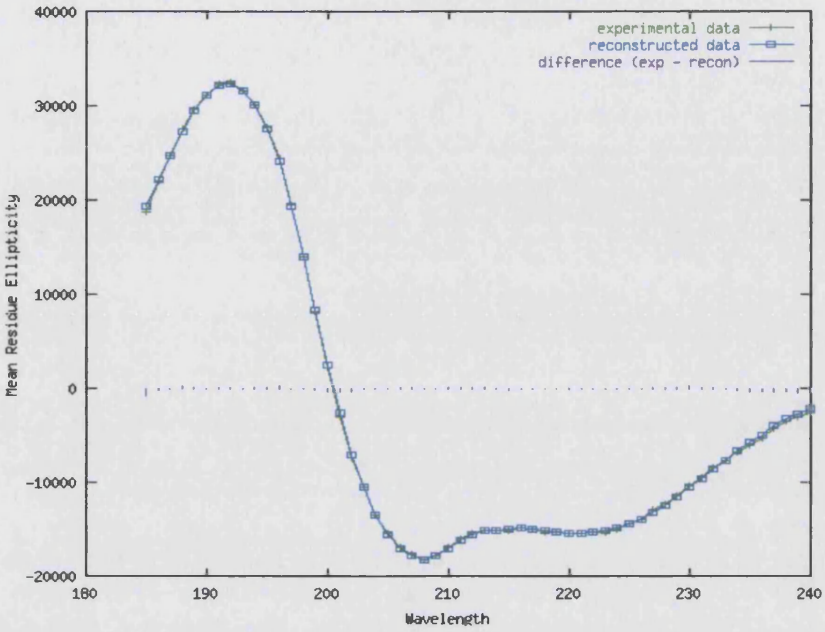


Fig. 3.9 Far-UV CD spectrum of Mns. Far-UV CD data were collected from Mns concentrated to 0.383mg/ml in 5mM phosphate buffer at pH 6.0, in a 0.02cm pathlength quartz cuvette. CD was performed by Sharon Kelly (University of Glasgow). The top panel shows the molar ellipticity, with a large positive maximum at 192nm and negative maxima at 208 and 220nm, indicating a largely α-helical structure. The bottom panel is the high tension voltage reading, and reflects the total absorbance from the sample. HTV remains below 550V to ~182nm, indicating the detector was not saturated and reliable data were collected.



Helix1	Helix2	Strand1	Strand2	Turns	Unordered	Total
0.36	0.22	0.04	0.04	0.16	0.18	1

Helix segments per 100 residues: 5.509	**Strand segments per 100 residues:** 1.770
Average helix length per segment: 10.541	**Average strand length per segment:** 4.052

Fig. 3.10 'Goodness of fit' CD spectra of Mns. Graphical output of far-UV CD data for Mns from Dichroweb (Whitmore & Wallace, 2004). Experimental data (green crosses) and calculated spectrum (blue boxes) derived from the calculated output secondary structure show very close agreement. The difference spectrum (purple lines) is very close to zero throughout the wavelength range, indicating the goodness of fit of the structural predictions. Mns is largely helical (58%), with only a small fraction of residues forming β-strands. Helix1 and Strand1 refer to 'regular' structures, and Helix2 and Strand2 refer to 'distorted' structures (Sreerama *et al.*, 1999).

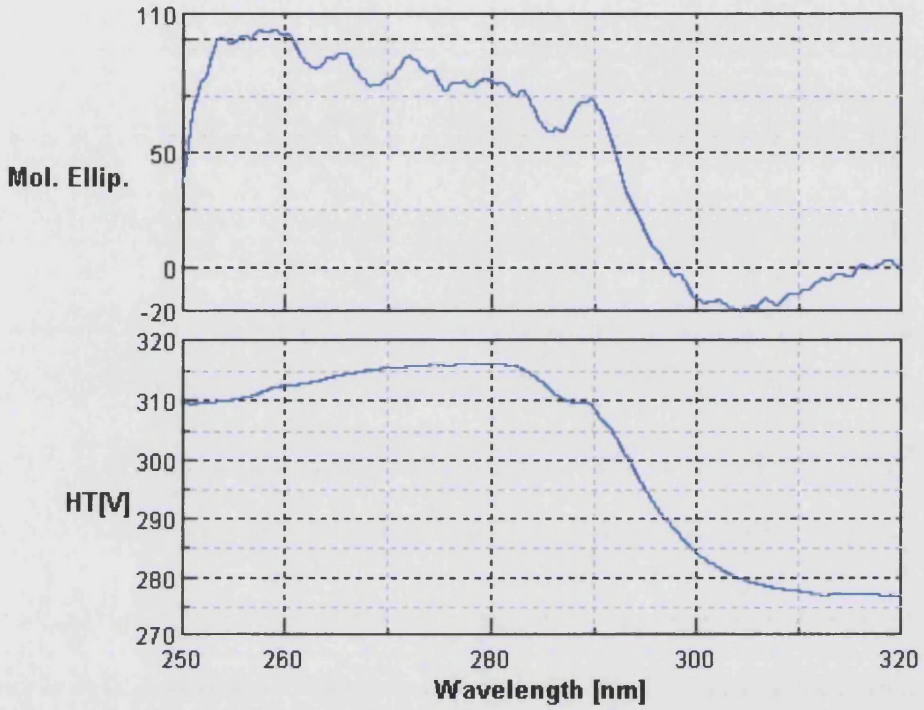


Fig. 3.11 Near-UV CD spectrum of Mns. Near-UV CD data were collected from a 0.5cm pathlength quartz cuvette containing Mns diluted to 0.773mg/ml.). The top panel shows a very weak signal was recorded in the near-UV wavelengths. The bottom panel shows the high tension voltage reading, which remains within the instruments boundaries for the collection of reliable data (below 550V). CD was performed by Sharon Kelly (University of Glasgow).

Fig. 3.12 Mns secondary structure predictions. **(A)** SSpro predicts Mns to have 59 residues forming strands (E, 43%), 7 residues forming helices (H, 5%) and the remainder (70/136 residues, 51%) forming coil (C) structures (i). The majority of the prediction is characterised by alternating strands and coils. Similar but not identical predictions were made by Jufo (ii) and PredictProtein (iii). **(B)** Mature CsgA protein sequence with structural elements predicted by SSpro. An alternating pattern of strands and coils is predicted, but is not fully consistent with a pattern of repeated units of strands (underlined) separated by 4-residue turns identified by Hammar *et al.* (1996).

3.3.5 Thermal Analysis

A purified sample of Mns was analysed by differential scanning calorimetry. The sample was heated at 15°C/min to 95°C, allowed to cool to room temperature and re-scanned. Energy changes within the system were recorded and normalised to allow for baseline correction. The normalised data (**Fig. 3.13**) show that enthalpy within the system was constant until the applied temperature reached 60°C. Above this temperature Mns begins to unfold, and reaches a midpoint (T_m) at 77.4°C, the point at which 50% of the protein molecules are unfolded. The thermogram then dropped toward the baseline and the protein was fully unfolded at 95°C.

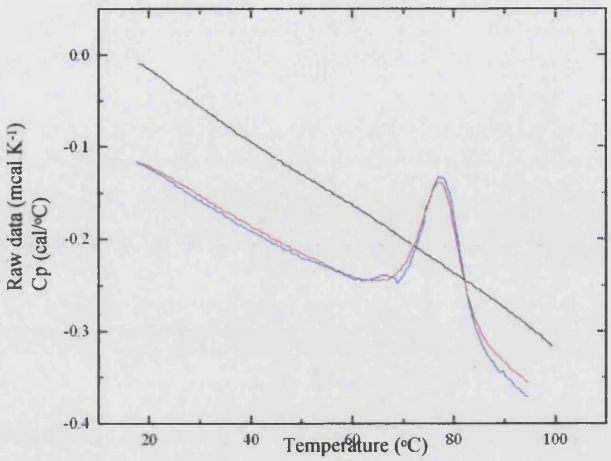
Upon re-scanning, Mns produced an almost identical thermogram, with the exception of a small shoulder between 60°C and 70°C, indicative of misfolding during the cooling process. There was no indication of aggregation following thermal denaturation, as judged by the shape of the thermogram (**Fig. 3.4**).

3.3.6 Dynamic Light Scattering

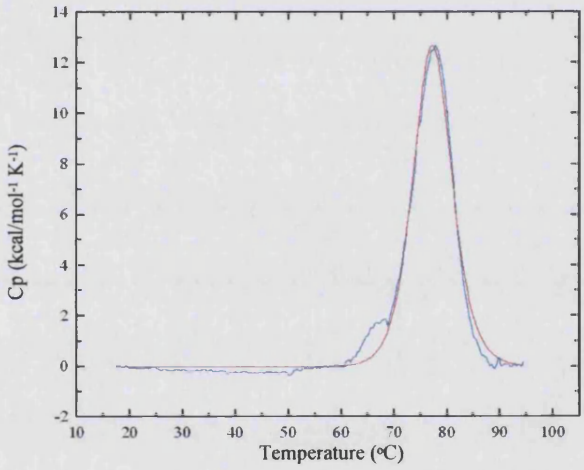
A purified sample of recombinant Mns was prepared to a concentration of 1.0mg/ml for NMR. The sample aggregated during NMR and analysis by DLS revealed the sample was polydisperse with large aggregates (**Fig. 3.14**). The DLS Protein Utilities function (Nano-S instrument software, Malvern) estimated that the smallest of the species (5.27nm) had the hydrodynamic diameter of a globular protein of approximately 32kDa and could conceivably be a dimer of the 14.9kDa Mns.

The sample was then filtered through a 0.022μm filter syringe, which produced a small volume (~20μl) of dilute protein that was reanalysed. This sample was relatively monodisperse, with 88.8% of particles existing in the same species of 5.63nm hydrodynamic diameter. A small fraction (11.2%) of the particles existed as an extremely large aggregate of 686nm in diameter (**Fig. 3.15**).

(A)



(B)



$\chi^2 = 135721$	$T_m = 77.43 +/- 0.0325$
$\Delta H = 1.24E5 +/- 923$	$\Delta H_v = 1E5 +/- 926$

Fig. 3.13 Differential scanning calorimetry of Mns. **(A)** Raw data from differential scanning calorimetry of Mns, showing a change in enthalpy above 60°C as the protein unfolds. The red line represents the initial analysis, and the blue line after re-equilibrating to the original start temperature and repeating the analysis. The black line shows the thermal profile of the buffer. **(B)** Normalised DSC thermal transition curve, adjusted to account for the thermal profile of the buffer and normalised with respect to protein concentration. Energy uptake by Mns reached a peak (T_m) at 77.4°C, representing the temperature at which 50% of the protein molecules are unfolded. This was almost identical after cooling the sample and repeating (blue line). The calorimetric enthalpy (ΔH) and Van't Hoff enthalpy (ΔH_v) are tabulated.

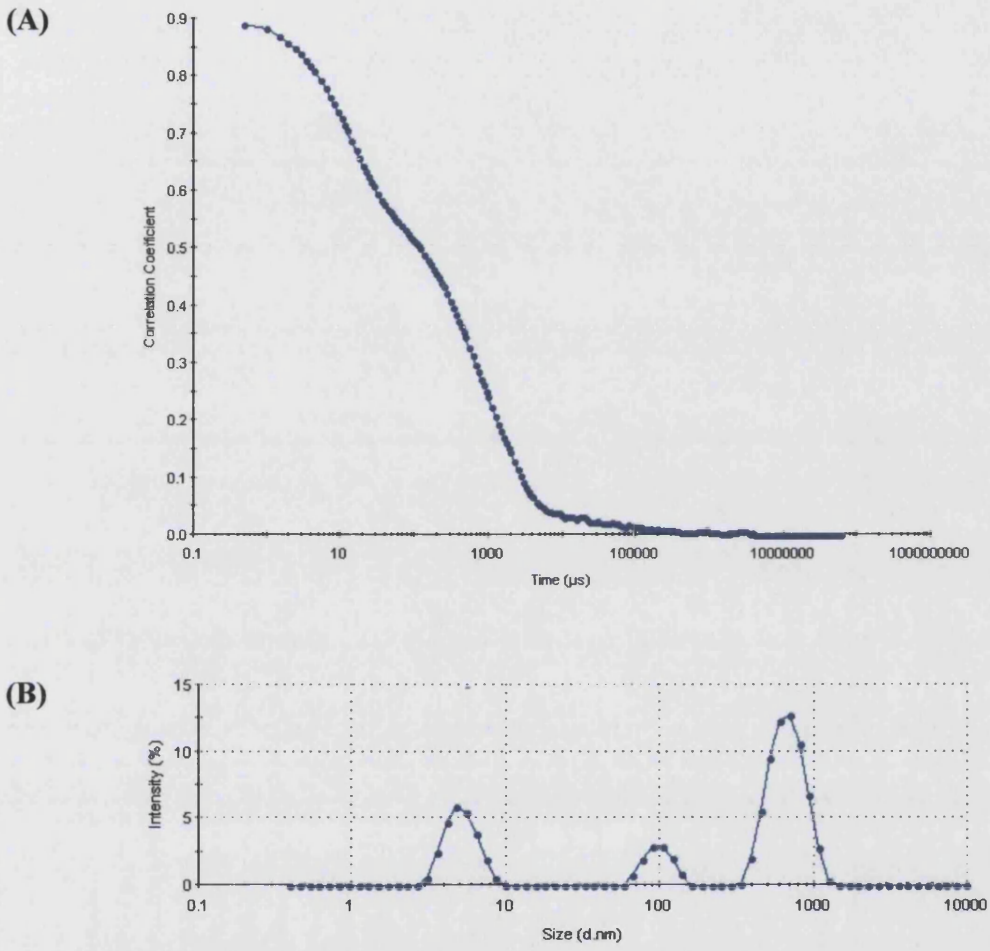


Fig. 3.14 Dynamic light scattering of Mns. Correlogram **(A)** and intensity **(B)** measures of concentrated Mns prepared in 5mM phosphate pH6.0, 150mM NaCl for NMR. The correlogram shows the time at which the DLS correlation starts to decay and indicates a large mean diameter of particles in the solution, and shoulders in the curve indicate the sample is polydisperse. The intensity reading shows that three species are present, with hydrodynamic diameters of 698nm (63% intensity), 5.27nm (25.6%) and 101nm (11.5%), which are estimated as 3000MDa, 32.5kDa and 32.5MDa respectively. This suggests that Mns exists as a dimer and in large aggregates of >1000 units.

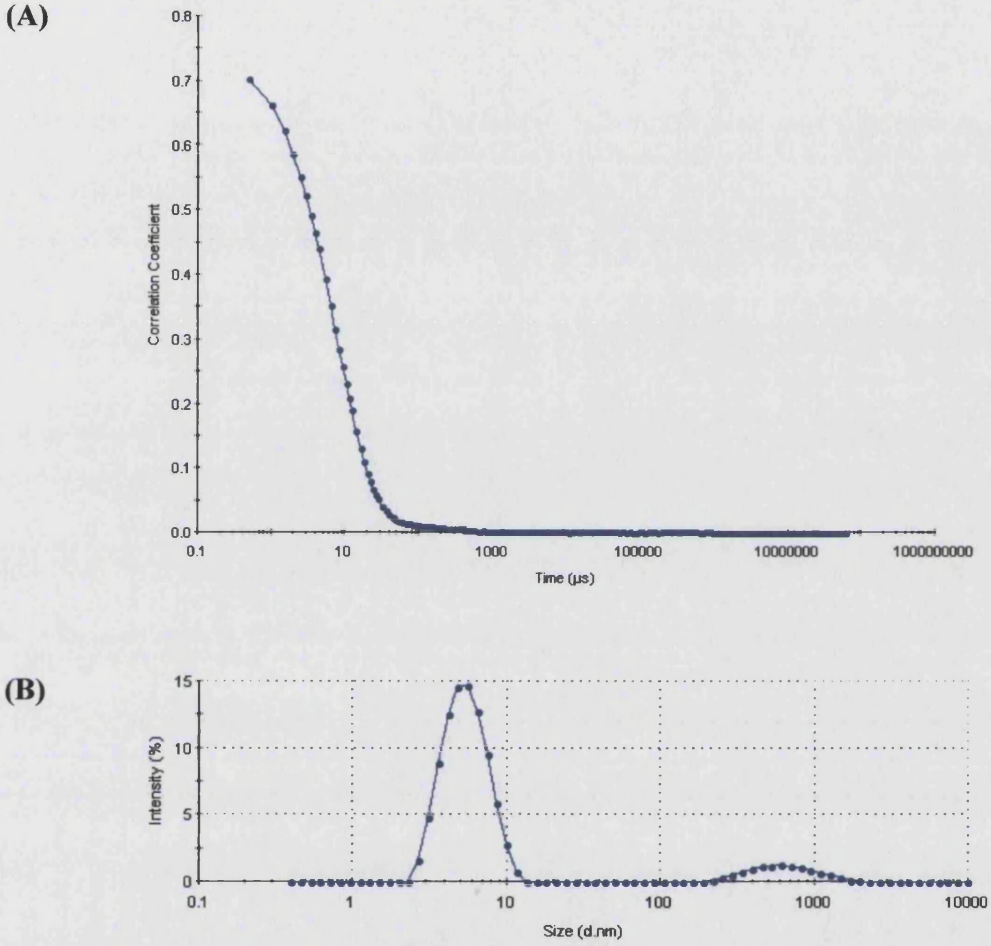


Fig. 3.15 Dynamic light scattering of Mns. Concentrated Mns was filtered with a 0.022µm filter and re-analysed by DLS. The steep angle of the correlogram suggests the particles are small in diameter **(A)**. The intensity measurement shows that 88.8% of the particles are of the same size, 5.63nm in hydrodynamic diameter **(B)**. This is the diameter expected of a globular protein weighing 37.9kDa. The protein also exists in a much larger aggregation, of estimated 2900MDa.

3.3.7 Analytical Gel Filtration

The suggestion that Mns forms a dimer is supported by analytical gel filtration. The recombinant protein was purified by ion exchange chromatography, concentrated and passed through a 75ml gel filtration column. Reference samples were then passed through the column. The Mns elution peak was at 77min after introducing to the column, and was very close to the elution peak of the 29kDa reference protein (**Fig. 3.16**). A Relative mobility calibration graph was constructed and the calibration trendline was used to calculate the size of Mns. An elution peak at 77min corresponds to an estimated size of 27.4kDa.

3.3.8 Congo Red Absorbance

Mns was incubated with Congo red to examine the possibility that the protein forms amyloid-like fibrils. No shift in absorbance was measured in the range 400nm-650nm in the presence of Mns compared to a no-protein negative control (**Fig. 3.17**). By contrast, a clear signal was obtained from the prion protein, which is known to form amyloid fibrils and cause the characteristic shift of absorbance in the presence of CR at 541nm (Chapman *et al*, 2002). These data suggest that purified Mns is not able to form amyloid fibrils.

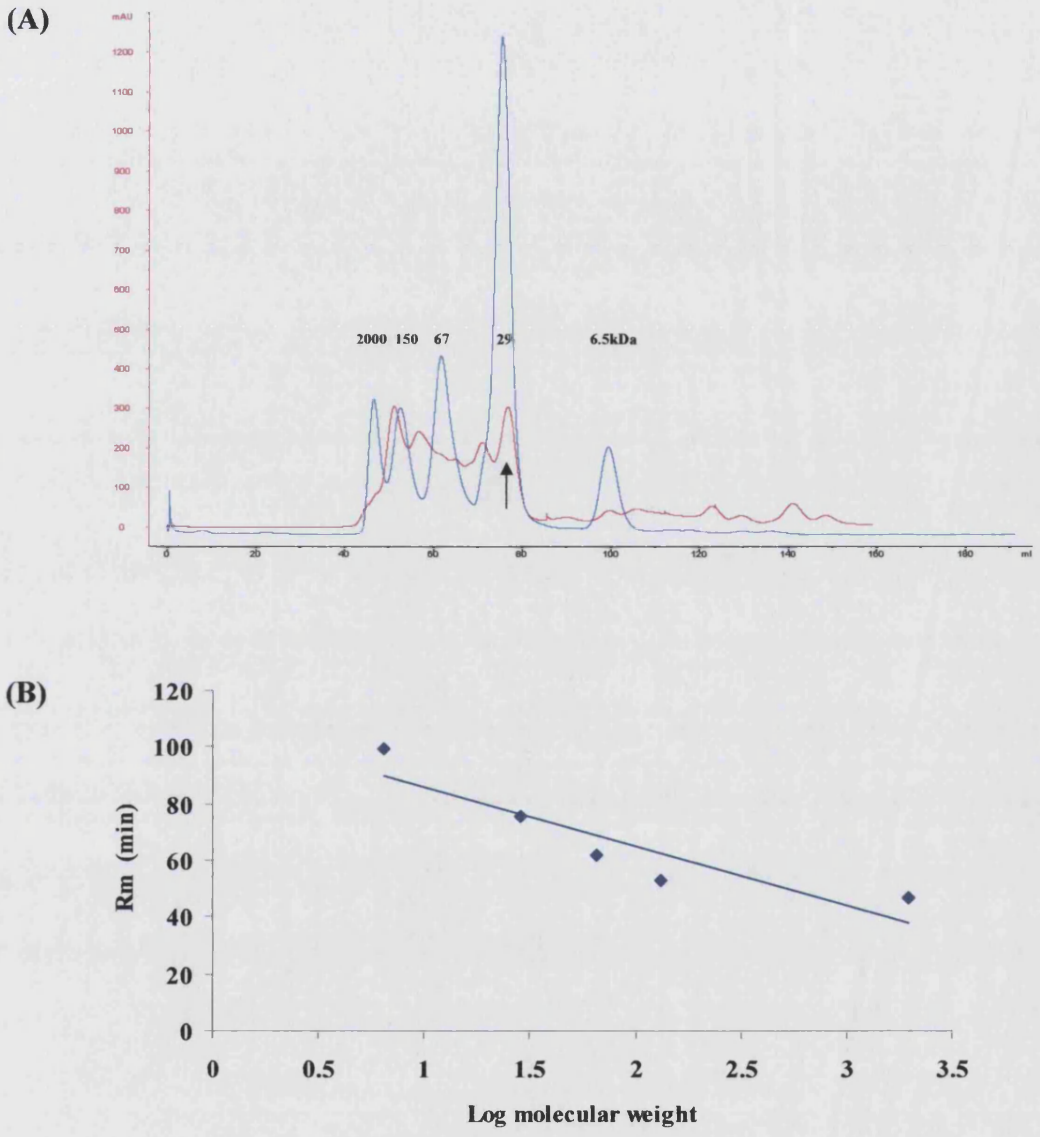


Fig. 3.16 Analytical gel filtration. **(A)** Partially purified Mns was filtered and passed through a Superdex 75 gel filtration column. The chromatogram shows Mns elution (red line) overlaid with an elution curve of reference proteins (blue line, with protein sizes as indicated- aprotinin (6.5kDa), carbonic anhydrase (29kDa), bovine albumin (66kDa), alcohol dehydrogenase (150kDa) and blue dextran (2000kDa)). Mns elutes close to the 29kDa peak, suggestive of a 14.9-14.9kDa dimer. **(B)** Relative mobility (Rm) calibration graph prepared using the elution times of the reference proteins. The peak of Mns elution was at 77min, which can be estimated from the graph as a log molecular weight of 1.44, or 27.4 kDa.

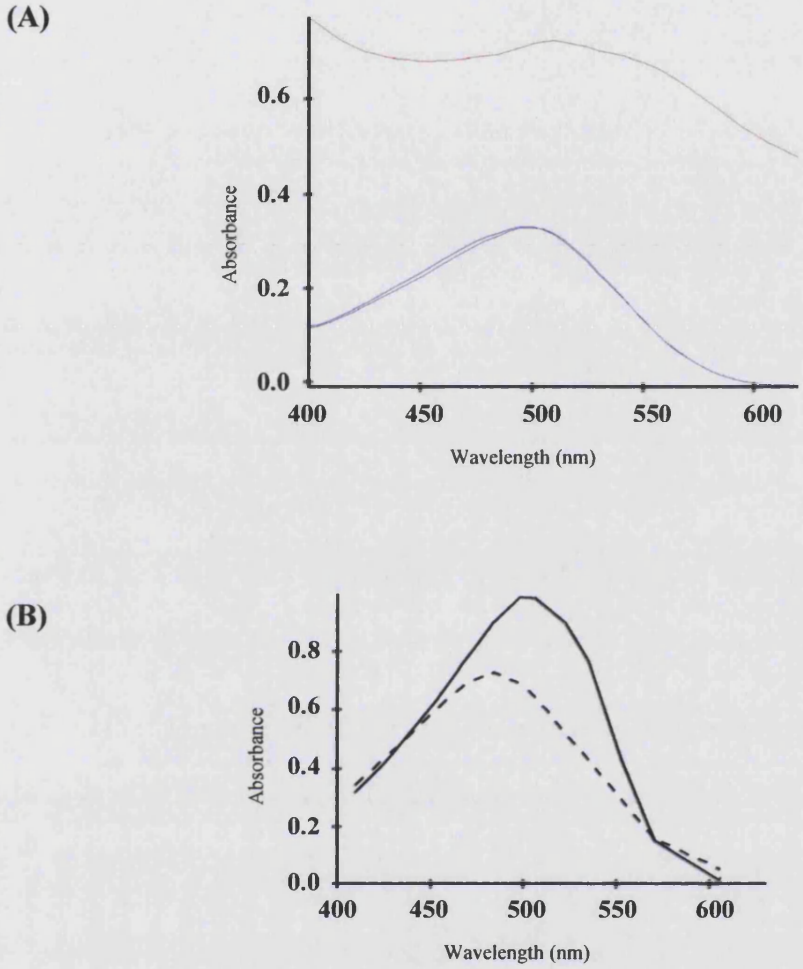


Fig. 3.17 Absorbance of Congo red by Mns. **(A)** A 10μM Congo red (CR) solution was mixed with 50μl recombinant Mns (4mg/ml) and made to a final volume of 1ml with 20mM HEPES buffer, pH7.5. The sample was scanned from 400nm to 650nm (blue) in a spectrophotometer. The absorbance profile did not differ from that of CR mixed with a buffer control (purple). The spectrum of 50μl of prion protein PrP shows an absorbance shift and is characteristic of amyloid fibrils. **(B)** Red spectral shift induced by amyloid fibrils. Curli fibrils mixed with CR solution (solid line) produced a shift in absorbance compared to CR in buffer alone (broken line). The characteristic shift in absorbance for amyloid fibres is 541nm. (Adapted from Chapman *et al.*, 2002).

3.4 DISCUSSION

3.4.1 Nuclear Magnetic Resonance

Mns is close to 15kDa in molecular weight and can be readily purified as a recombinant protein. The protein is also soluble, making it amenable to study by nuclear magnetic resonance. By transforming the *mns*-pBAD30 vector into BL21 Gold *E. coli* we were able to grow the cells in minimal media and produce protein with labelled carbon and nitrogen atoms, a necessary step for the collection of ^{15}N- and ^{13}C-NMR data and assignment of NMR signals to specific atoms in specific amino acids.

Unfortunately, under the concentrated conditions required for NMR the protein aggregated over periods ranging from less than one day up to fourteen days for the first unlabelled sample, and did not allow complete data sets to be collected. Different purification techniques were used, and the protein was exchanged into different buffers in an attempt to reduce the aggregation rate. Such approaches have been used elsewhere and successfully improved monodispersity (Nominé *et al.*, 2001). Furthermore, we purified Mns proteins with altered amino acid sequences in an attempt to reduce autoaggregation but were unsuccessful.

The data collected by NMR were insufficient to provide any detailed insight into the structure of Mns. The excellent chemical shift dispersion observed in the 1D and ^{1}H-^{15}N HSQC spectra is consistent with the presence of a substantial amount of β-secondary structure. Aggregation of the protein resulted in poor signal intensity, with broad peaks that were difficult to distinguish from the noise.

The development of cryogenic probes has increased the sensitivity of NMR spectrometers significantly, reducing the signal-to-noise ratio by an average of four-fold (Spraul *et al.*, 2003). For the collection of data on a protein such as Mns this has two implications. Firstly, this sensitivity allows a four-fold lower detection limit, meaning the protein can be of a lower concentration and therefore be less likely to aggregate. Secondly, the experimental time can be reduced by a factor of 16. If aggregation or degradation of the protein is a time dependent phenomenon this has

obvious advantages. A sample of Mns was prepared for spectroscopy with an 800MHz spectrometer fitted with a cryoprobe, but despite the lower concentration used the protein aggregated and the spectra collected were no better than those collected previously.

3.4.2 Crystallisation

Samples submitted to the Hauptman Woodward Institute were screened using the microbatch-under-oil method (Luft *et al.*, 2003). Small crystals were obtained from a number of conditions but could not be reproduced, despite varying the protein concentration and concentration of the different components. Interestingly, several conditions from which crystals were obtained contained metals, including zinc, manganese and sodium, inviting us to believe the protein may interact with metals.

A sample prepared for NMR was treated with EDTA and suspended in a Chelex-treated buffer to remove all metals. Whether this caused any conformational change is unknown, but a ^{1}H-^{15}N HSQC initially showed about 40% fewer peaks than samples prepared in a standard fashion. This result was not reproducible and may be worth revisiting in the future. The protein showed a similar pattern of aggregation to untreated samples.

Crystals that form under batch conditions may not readily form crystals by vapour diffusion, and some care is needed when interconverting between the two methods (Weincek, 1999). This is because supersaturation is met immediately when the components of the drop are mixed and are not reached after equilibration within a well. Additionally, the difference may in part be due to the oil used to protect the drops from evaporation, which can affect nucleation, crystal growth and stability (Chayen, 1999). Interestingly, D'Arcy *et al.* (1996) have shown that using a mixture of oils causes the drop volume to shrink, resulting in a gradual concentration of protein and precipitant rather like the process that occurs in the vapour diffusion method.

The largest crystals obtained for Mns came from 8% PEG 8000, 0.1M Tris HCl pH 8.5, conditions that were identified from an initial screen kit and grew after seeding microcrystals into a fresh drop of the same precipitant and protein conditions. One such crystal was prepared for X-ray crystallography but was revealed to be a salt crystal. Empirical observations suggest that proteins that exist in monodisperse solutions in a single aggregation state are likely to crystallise and that polydisperse aggregates fail to yield good crystals (Ferré-D'Amaré & Burley, 1997). The aggregation we observed by NMR and dynamic light scattering may have hindered crystallisation, and been a particular problem at the high concentrations needed.

3.4.3 Circular Dichroism Spectroscopy

The circular dichroism of a protein is a direct reflection of its secondary structure. The most successful way to correlate the two is to obtain the CD spectra of polypeptides in known structural conformations, and use these as basis spectra when making predictions about structurally uncharacterised proteins (Hennessey & Johnson, 1981). Here, Mns was analysed by both far- and near-UV CD spectroscopy, with the aim of making an initial characterisation of the structural features of the protein. Visual examination of the far-UV spectrum shows that Mns doesn't form a random coil but contains defined structural elements. The protein appears to be largely α-helical, with a strong positive maximum at 192nm, and negative maxima at 208 and 222nm. However, the 222nm maximum is low in magnitude, and a small negative maximum exists at 215nm, which is suggestive of β-strand components (**Fig. 3. 10**).

3.4.3.1 Secondary Structure

For a more comprehensive analysis of secondary structure, the CD spectral data were analysed using the software CDSSTR (Oregon State University, Johnson, 1999). The algorithm is based on the variable selection method, which uses the spectral data of a fixed number of proteins chosen from a reference set. Here, CDSSTR used all combinations of 7 basis spectra from a reference set of 37. The secondary structures of all the proteins in this set have been determined by refined X-ray diffraction data with a resolution of at least 2.0Å (Johnson, 1999). Each protein in the reference set provides a different basis curve with unique maxima for comparison with the test protein. When performing the algorithm, these reference proteins are eliminated if

they (a) possess structural characteristics not seen in the test protein, or (b) do not have the structural characteristics the test protein possesses.

Flexibility in the system, comparing the test protein against 7 best-fitting spectra rather than a complete reference set, allows the two main criteria in secondary structure prediction to be met. Firstly, there should be no negative fractions of secondary structure (there can be no less than 0 α-helices, and no less than 0 β-strands etc.). Secondly, the sum of all fractions must be close to 1.0 (Johnson, 1999).

There may be many combinations of 7 basis spectra that successfully meet these requirements, each possibly giving a different prediction of the characteristics of the test protein. To select the analyses that most accurately predict the secondary structure, the spectrum of the test protein is fitted against all the reference set spectra. This uses the singular value decomposition (SVD) method without basis-set flexibility, and has been shown to give the best prediction of α-helix content (Greenfield, 1996). The final result is a prediction whose potential accuracy is close to the variation in X-ray diffraction data. For Mns, the prediction is of an ordered secondary structure, 58% of which is α-helical. Only 8% of residues form β-strands, 16% form turns, and the remainder are disordered (**Fig. 3.10**).

The normalised root mean standard deviation (NRMSD) is a goodness-of-fit parameter, and represents the extent of agreement between the measured CD spectrum and the predicted structure; NRMSD values of control samples, for which X-ray diffraction data are available, indicate an excellent agreement between the predicted secondary structure and X-ray structure when less than or equal to 1.0. NRMSD values greater than 2.0 indicate that the calculated and actual structure do not bear likeness to one another (Mao *et al.*, 1982). The obtained value for Mns was 0.011, and suggests the predictions of structural fractions made from the basis set spectra are very consistent with the actual structure. However, although a low NRMSD is necessary it is not sufficient to confirm the accuracy of a prediction, because the NRMSD is ultimately dependent upon how closely the reference proteins share structural similarity to Mns (Mao *et al.*, 1982).

A single structure leads to a unique spectrum but the inverse is not always true. Glättli *et al.* (2002) have found, with non-natural oligomers of β-peptides, that a given CD pattern can be generated by spatially different structures. This complicates the interpretation of CD spectra and may lead to false conclusions, further implying that the conformation of a peptide cannot be unambiguously derived from CD measurements.

Analysis of CD data is further complicated by the factors that contribute to the signal other than secondary structures. For instance, the length of α-helices and twists in β-sheets also affect the signal, as do aromatic and sulphur-containing amino acids that absorb light in the far-UV region (Krittanai & Johnson, 1997, Johnson, 1999). Algorithms have been developed to overcome these factors, including the programme CDSSTR that has been used in the present study, ensuring the results obtained for Mns are as accurate as the state of the art allows.

3.4.3.2 Tertiary Structure

To gain insight into tertiary structural characteristics, CD data may be collected from the near-UV range, which is extremely sensitive to conformational changes such as those induced by changes to pH, temperature or the binding of ligands. The near-UV CD of a protein arises largely from absorbance by the aromatic side chains (Kelly & Price, 2000). Mns contains one phenylalanine, three tryptophan and eight tyrosine residues, which together comprise 8.8% of the total residues and explain the weak spectrum recorded for the protein. Disulphide bonds also absorb in the near-UV wavelengths and may contribute to the CD spectrum near to 260nm, but Mns contains just one cysteine residue and therefore is not expected to produce such interactions.

In the region 400-500nm the presence of a small spectrum would be indicative of a bound flavin (Kelly & Price, 2000). Flavins are cofactors for many enzymes and are known to give a yellow colouration to proteins (Zhu *et al.*, 2006). When Mns was highly concentrated, a yellow colour was observable, suggesting it may bind to flavin. CD data for Mns was not collected above 320nm, but no detectable absorption spectrum was observed in this region when the protein was analysed with scanning spectrophotometry (data not shown) as has been reported elsewhere for bound flavins (Schmitz, 1997).

Near-UV CD is sensitive to changes in tertiary structure because, as the protein changes conformation, the environment that surrounds the aromatic amino acids and disulphide bonds may change, and this is associated with changes in the CD spectrum. Similarly, the reduction or release of bound flavin groups will be reflected in changes to the spectrum (Kelly & Price, 2000). Near-UV data is therefore most useful when collected from proteins in different experimental conditions; the current understanding of CD in this region is not sufficient to make predictions from a single spectrum.

Some degree of tertiary structural prediction is available from far-UV data. Manavalan & Johnson (1983) showed that key features of CD spectra in this region allow identification of proteins belonging to each of four classes previously described by Levitt & Chothia (1976): all-α (predominantly α-helical), all-β (predominantly β-sheet), α+β (separate α-helix- and β-sheet-rich regions) and α/β (intermixed segments of α-helix and β-sheet). The first feature is the wavelength at which the spectrum crosses from positive to negative. For Mns, data was not collected below 180nm, but by extrapolation of the spectrum to 170nm one might estimate that the crossover occurs at a wavelength above 172nm; an indication that Mns does not belong to the all-α group. The magnitude of the 208nm negative maxima relative to the 220 maxima is a further indicator, and places Mns among the type α+β proteins.

Type α+β proteins consist of all-α and all-β regions within the same polypeptide chain; the α-helix and β-strand secondary structure segments do not mix but tend to segregate along the chain. Other members include cytochrome b_5, hen lysozyme and ribonuclease, which have a cluster of helices at one or both ends of the β-sheet, which is often built up from strands in antiparallel associations. In larger type α+β proteins stable domains can be identified in which one domain is largely helical and the other mainly consists of β-sheets (Levitt & Chothia, 1976).

For Mns, the number and average length of the α-helix and β-strand segments have been predicted. These estimates derive from averages of the frequency of distorted and regular α-helices, which can be distinguished by CD. When secondary structures have conformational angles that deviate from ideal angles in geometrically pure helices the deviation causes distortions such as twists and bends. Such geometric variations are believed to influence the structure of a CD spectrum (Sreerama *et al.*, 1999). Using an average number of four distorted residues per α-helix and two distorted residues per β-strand, and integrating these averages with fractions of distorted and regular structures estimated by CD analyses, it is possible to make estimates of the frequencies and length of these secondary structural elements.

For Mns, 5.51 helix segments per 100 residues were predicted, or 7.49 helices in the 136-residue protein, and each helix was estimated to be an average 10.54 residues in length. CD has been used to show the implications of helix length; as length increases, so too does stability. This length-dependent effect on conformational stability arises because the energetic cost of initiating helix formation is much greater than that of adding additional residues to the helix (helix propagation) (Chakrabartty & Baldwin, 1995). Additionally, the effect of end-residue fraying proportionally decreases as helix length increases, favouring more extended structures (Rohl *et al.*, 1992). Zimm *et al.* (1959), however, noted that the helices of globular proteins must be quite short, and limited in length by the dimensions of the globular molecule if not otherwise. On average, the helices of Mns are close in length to the average of 9.24 taken from a reference set of proteins from the Protein Data Bank (Sreerama *et al.*, 1999)

The stability of β-strands is also length dependent, and is two-dimensional, occurring along the strand direction and perpendicular to this direction. Stanger *et al.* (2001) have reported increases in conformational stability from 5 to 7 residues per strand in antiparallel β-sheets, but no further increases with extension to 9 residues, and suggested that as strand length grows, so does a propensity within each strand to form an α-helix. In Mns, few strand segments were predicted; 2.41 strands (1.77 per 100 residues). The average length was 4.05 residues per strand, and is again close to the average (5.02 residues) from a range of proteins (Sreerama *et al.*, 1999).

3.4.3.3 Reliability of Circular Dichroism Data

Circular dichroism is a low-resolution technique. The data presented here offer estimates of the type of secondary structures the protein forms, as well as an indication of the frequency and length of these segments, but no information regarding the specific residues that form these structures or their spatial distribution. The recorded spectrum closely matches a spectrum calculated from reference spectra, with a normalised root mean standard deviation of 0.011; as far as the reference spectra reflect the structures present in Mns, the predictions are highly reliable.

Experimental reliability depends on accurate measurement of the test protein concentration, as well as instrument calibration and knowledge of the cell pathlength (Miles *et al.*, 2005). It is also necessary to keep the total absorbance of the measurement within specified bounds in order to avoid excess noise. When a radiation source is used in CD, the intensity of an emitted signal drops significantly towards the lowest wavelengths in its range: towards 250nm for near-UV and towards 180nm for far-UV CD spectra. The detector has to increase its sensitivity in order to collect low intensity signals, and this requires an increase in high tension voltage; there is a maximum HTV at which the instrument can accurately record the transmitted radiation, and towards this maximum data can be unreliable (Kelly *et al.*, 2005). To meet this criterion for Mns, the protein concentration and cell pathlength were adjusted, and the data collected is reliable within the wavelengths 320 to 250nm and 260 to ~185nm.

CD spectroscopy is most effectively used to complement more detailed structural analysis, and was used in the present study to complement NMR spectra, with the potential to prepare samples in different buffers or in the presence of possible ligands to reveal the structural changes they induce.

3.4.4 Analytical Gel Filtration

Gel filtration is commonly used as a final purifying step in the preparation of proteins for X-ray crystallography. Because the technique separates particles according to size, it can also be used as an analytical procedure to make an estimate of protein size in non-denaturing conditions. The Mns elution peak corresponded to that of a 27kDa protein. This suggests that the smallest unit detectable in a purified sample was a dimer, and these data were supported by dynamic light scattering analysis.

3.4.5 Dynamic Light Scattering

DLS analysis clearly showed the propensity of Mns to aggregate. Only by filtration were we able to obtain a solution without major aggregation. The estimated hydrodynamic diameter of the smallest particle in the suspension was in the range 5.3-5.6nm. This estimate, based on the velocity of particles moving in fluid, provides an estimate of particle mass: 32.5-37.9kDa. It is not inconceivable that these particles are dimers of Mns. Other particles in the sample, still present after filtration and not removed by ultracentrifugation (data not shown) were up to approximately 100-fold greater in diameter and likely acted as nuclei to more aggregation (Fändrich, 2007).

3.4.6 Prediction of Aggregation Propensity

In high concentrations almost any protein can be induced to form aggregates. Most will be disordered, and formed from non-specific hydrophobic contacts between misfolded proteins. Alternatively, an aggregate may be enriched in cross-β structures and can form if an amino acid sequence has a strong propensity to be in a β-extended conformation. Only a small subset will form ordered structures, rich in β-sheets and known as amyloid fibres. Fibres are almost crystalline in their stability and their amino acid composition is position specific. Several peptide sequences have been identified that are strongly amyloidogenic, including GNNQQNY, KFFEAAAKKFFE, YTIAALLSPYS and the hexapeptide STVIIE (reviewed by Rousseau *et al.*, 2006b). None of these sequences, in whole or in part, exists in Mns.

Two algorithms are available for the prediction of the propensity of a protein to aggregate. The first, TANGO, is based on the physico-chemical principles of β-sheet formation and experimental data of proteins known to form self assemblies (Fernandez-Escamilla, 2004). The second, Zyggregate, predicts aggregation-susceptible regions, and like TANGO is based on the identification of the propensity of individual amino acids to cause aggregation (Pawar *et al.*, 2005).

A TANGO plot of β-aggregation for Mns identified two regions that are predicted to have a high intrinsic propensity to aggregate and are likely to promote the aggregation of the entire sequence (**Fig. 2.24**). If at least five consecutive residues populate the β-aggregated conformation they are considered to have some aggregation tendency. In Mns, the two regions with aggregation tendency were of 6 and 9 residues in length. The average length, identified from thousands of proteins from all kingdoms of life, is also 9 residues (Rousseau *et al.*, 2006a).

β-aggregation in TANGO should not be confused with β-sheet propensity, and peptides with a high β-sheet propensity will not necessarily show aggregation (Fernandez-Escamilla *et al.*, 2004). In the case of Mns, the two regions predicted to have a high β-aggregation propensity were also identified to be involved in β-strands (**Fig. 3.12**). Finally, it should be pointed out that TANGO cannot be used to identify potentially amyloidogenic peptides because it does not capture the specific sequence contributions that appear to be important in fibre formation.

Aggregation of proteins is minimised by reducing the amount of strongly aggregating sequences, but the properties that favour aggregation are sometimes necessary to the structure of a globular protein. To contain aggregation, a universal mechanism appears to be the enrichment of 'gatekeeper' residues at the edge of aggregating sequences. These residues counteract the aggregation propensity of sensitive sequences and minimise misfolding as much as possible. Glycine and proline residues have structure-breaking properties, and the charged sidechains of arginine and lysine are repulsive in close packed aggregates and are effective at opposing aggregation (Rousseau *et al.*, 2006b). Consistently, the predicted aggregation sequences of Mns are flanked by such residues, which potentially act as gatekeeper residues in curtailing the effect (**Fig. 2.24**). Lastly, analyses of aggregation propensity

have shown that a protein is equally likely to aggregate whether it belongs to any of the structural groups (all-α, all-β, α+β or α/β) described by Levitt and Chothia (1976, Rousseau *et al.*, 2006a). Therefore, the classification by circular dichroism of Mns as an α+β protein with a low β-strand content should not be considered an impediment to its aggregation propensity.

3.4.7 Implications of Aggregation

The aggregation of Mns may not be a consequence of misfolding or an artefact of preparation. Misfolded proteins generally precipitate out of solution, and although the possibility exists that Mns may eventually have precipitated, the protein appeared to remain soluble even after extensive aggregation. As discussed in Chapter 2, the protein is able to adhere to surfaces and is implicated in aggregating nematodes, and the self-association phenomenon we observed may be crucial to these activities. Moreover, it may be significant that there is an apparent maturation of *P. asymbiotica* supernatant. Supernatants incubated at 4°C for an extended period such as one week were better able to cause the nematode aggregation phenotype than supernatants prepared freshly from *P. asymbiotica* cultures (unpublished observation).

In *E. coli,* curli fibres are produced that facilitate binding to abiotic surfaces and the formation of biofilms, and these are assembled from highly aggregated curlin subunits. The major structural component of curli is CsgA, which efficiently polymerises *in vivo* in the presence of CsgB and CsgEFG. Interestingly recombinant CsgA has been purified and may exist *in vitro* as unaggregated curlin, but prolonged incubation at 4°C results in aggregation in the absence of CsgB and the formation of fibres similar to those produced by wild-type bacteria (Chapman *et al.*, 2002). The conversion to fibrils is associated with an increase in β-sheet composition, forming the cross-β structure that defines amyloid fibrils (Sipe & Cohen, 2000).

CD analysis of Mns has suggested the protein has a low β-strand content, but changes in secondary structure have been shown to occur as proteins aggregate, such as the conversion of α-helices in the prion protein to β-sheets in the scrapie form (Pan *et al.*, 1993). Therefore the possibility exists that Mns may form β-stranded structures and assemble into polymeric forms. Interestingly, secondary structural prediction programmes estimate a higher strand composition than was obtained by CD, and

NMR data also suggested a substantial β-sheet structure. The SSpro program is the most accurate server available and achieves a three-state (helix, strand or other) prediction accuracy of >78%. SSpro predicted that 43% of Mns residues form strands and just 5% of the protein is helical. Relatively high levels of β-strand composition were also found with Jufo (38%) and PredictProtein (55%).

The predicted strand and coil elements of the secondary structure of Mns occur alternately throughout the sequence, with little interruption from helices. A similar, but somewhat more regular pattern has been described for the major and minor curlin subunits (Hammar *et al.*, 1996). CsgA and CsgB are considered homologs, but share only 30% identity. Their similarity exists through common secondary structural motifs; repeated units of β-strand (7-8 residues), turn (4 residues), β-strand (7 residues), turn (4 residues) (Hammar *et al.*, 1996). It may be significant that these units were not accurately predicted using the SSpro programme (**Fig. 3.12**).

As a consequence of the tendency of Mns to aggregate, *Photorhabdus* may require machinery that prevents polymerisation within the cell, or enhances this process at the cell surface. Such nucleation machinery has been described elsewhere for *E. coli* (Hammar *et al.*, 1996).

3.4.8 Thermal Analysis

As an exploratory purification method, samples of sonicated *E. coli* expressing Mns were heated to 40°C, 45°C and 50°C for 10min and the insoluble proteins were pelleted by centrifugation. We observed that despite the precipitation of many *E. coli* proteins, recombinant Mns was still present in the soluble supernatant (data not shown). Although the method was not used for subsequent purifications, we found it of interest to analyse the apparently high thermal stability of Mns. DSC of Mns produced a raw-data thermogram that shows the difference in heat capacity of Mns compared to a reference solvent buffer, and can be divided into three phases: a pre-transition heat capacity, a denaturation transitional peak, and a post-transition heat capacity (**Fig. 3.13**).

At low temperatures the C_p of Mns follows that of the reference solvent. At approximately 65°C Mns enters a transitional state and begins to denature. This process absorbs energy and the heat capacity increases to the midpoint temperature (T_m = 77.43+/- 0.03°C). At midpoint, 50% of the protein molecules are unfolded and 50% remain folded, and the shape of the thermal transition peak for Mns is characteristic of many single-domain proteins; the unfolding process is a two-state reaction, with a single cooperative event leading to the denatured state. Non-cooperative and multi-state transitions are more frequently observed when large, multi-domain proteins unfold and produce intermediary forms (Creighton, 1993). Above the midpoint temperature, the heat capacity of Mns drops to the baseline at approximately 89°C, with the C_p of the unfolded polypeptide again following that of the reference solvent.

The thermal stability of Mns is high among bacterial proteins, which more commonly denature around 50-65°C (**Table 3.1**). At higher temperatures, ribosomal proteins unfold (~74°C for *E. coli*). This characterises the DSC profiles of whole-cell preparations and ultimately heat-inactivates the bacteria. Transitions of DNA and outer cell wall components follow as temperatures are further elevated (Lee & Kaletunc, 2002).

3.4.8.1 Refolding of Mns after Denaturation

Rescanning Mns after a period of cooling shows that the protein is able to refold after thermally induced denaturation. This refolding is in the absence of chaperones, which are required for the correct folding of some proteins (Ellis, 1987). The amino acid sequence of Mns therefore inherently possesses enough information to allow the formation of secondary and tertiary structures, which may have implications related to secretion of the protein; if Mns is secreted through a type III system, the internal diameter of the secretion needle is expected to accommodate only single α-helices or randomly coiled structures (Deane *et al.*, 2006).

Organism	Protein	Temperature	Reference
Gram negative			
Borrelia bergdorferi	OspA	67.8°C	Koide *et al.*, 1999
Desulfovibrio gigas	Adenylate kinase	45°C	Gavel *et al.*, 2004
Escherichia coli	FhuA	74.4°C	Bonhivers *et al.*, 2001
Escherichia coli	GGBP	53°C	Piszczek *et al.*, 2004
Escherichia coli	MBP	64.9°C	Ganesh *et al.*, 1997
Escherichia coli	NAD+-DNA ligase	54.1°C	Georlette *et al.*, 2003
Escherichia coli	RNase H	66°C	Robic *et al.*, 2003
Photorhabdus luminescens	Mns	77°C	Present study
Pseudomonas diminuta	phosphotriesterase	~58°C	Rochu *et al.*, 2002
Shigella flexneri	MxiH needle	42°C	Darboe *et al.*, 2006
Gram positive			
Bacillus subtilis	Adenylate kinase	47.6°C	Bae *et al.*, 2004
Cellulomonas fimi	Cellulase CenC	55°C	Creagh *et al.*, 1998
Clostridium botulinum	BoNT A	50.5°C	Encinar *et al.*, 1998
Clostridium botulinum	BoNT E	51.1°C	Encinar *et al.*, 1998
Streptomyces hastedii	Xylansae Xys1	64°C	Ruiz-Arribas *et al.*, 1998
Psychrophile			
Bacillus globisporus	Adenylate kinase	43.3°C	Bae *et al.*, 2004
Bacillus sp TA41	Subtilisin	50°C	Davail *et al.*, 1994
Methanococcoides burtonii	Elongation Factor-2	37.4°C	Thomas *et al.*, 2000
Psychrobacter immobilis	β-lactamase	~50°C	Feller *et al.*, 1995
Thermophile			
Bacillus stearothermophilus	Adenylate kinase	>55°C	Bae *et al.*, 2004
Geobacillus stearothermophilus	α-Glucuronidase	73.4°C	Shallom *et al.*, 2004
Geobacillus sp TW1	Lipase	>90°C	Li & Zhang, 2005
Methanobacterium formicicum	Histone HfoB	76.5°C	Li *et al.*, 1998
Methanosarcina thermophila	Elongation Factor-2	49.7°C	Thomas *et al.*, 2000
Sulpholobus acidocaldarius	Sac7d	90.7°C	McCrary *et al.*, 1996
Thermus thermophilus HB27	Laccase	92°C	Miyazaki, 2005
Thermus scotoductus	NAD+-DNA ligase	95.6°C	Georlette *et al.*, 2003

Table 3.1 Melting temperatures of bacterial and archeal proteins. Proteins from a range of organisms were randomly selected from the available literature and provide a context for the melting temperature of Mns, which has a higher thermal stability than the other Gram-negative bacterial proteins identified.

The process of protein folding is not well understood, but likely begins with local interactions induced as the protein is exposed to non-denaturing conditions. These interactions lead to the formation of secondary structure. At this stage, there is little cooperativity between secondary structural forms, and few of the interactions will remain as contacts in the native protein.

Folding intermediates possessing largely native-like secondary structures will become further stabilised as the hydrophobic residues collapse into a core, creating a well-defined and compact structure with hydrophilic residues exposed at the surface. The polypeptide sequence tends to favour the destabilisation of incorrectly folded proteins such that the majority can reach the correct fold. Kinetically trapped intermediates have also been observed, which are unable to refold into productive intermediates and native proteins (reviewed by Miranker & Dobson, 1996).

In the case of Mns there are no detectable intermediates, and the folding process appears to proceed via a 2-state, monophasic transition. It is interesting to note that there was no apparent aggregation at the highest temperatures. Aggregation of unfolded polypeptides is indicated by distortion of the DSC thermogram and a C_p drop below the reference baseline due to the exothermic nature of the phenomenon. Aggregation is a kinetic process, promoted by higher temperatures and concentrations, and is irreversible; aggregated protein will not refold when cooled (Cooper, 2001).

In the absence of polypeptide aggregation, improper refolding or chemical changes induced by high temperatures may cause changes in the transition curve of a protein upon rescanning. Chemical changes may include oxidation of sulphur residues, deamidation of glutamine and asparagines, or racemisation about the Cα atom (Creighton, 1993). The occurrence of one or more of these processes is likely to have caused the minor shoulder before the major transition peak observed for Mns (**Fig. 3.13**). We currently have no direct assay to confirm that Mns is biologically active after refolding, or indeed if the protein is biologically active at the high temperatures it withstands before unfolding.

3.4.8.2 Enthalpy of Denaturation

In addition to the high temperature of denaturation, DSC of Mns revealed the enthalpy of denaturation (ΔH). This parameter is estimated in two independent ways. Firstly, the ΔH_{cal} is the area under the thermal transitional peak and is dependent upon the amount of protein in the sample. For Mns this was $1.25x10^5$ +/-923 cal mol^{-1} and is an absolute measure of the heat energy taken up by the denaturation process.

The second measure, the van't Hoff enthalpy, is based on the shape of the transitional peak and the fraction of protein molecules under it that are folded or unfolded. A sharp peak gives a higher estimate of ΔH_{VH} and is indicative of a highly cooperative transition from the native to denatured state. The ΔH_{VH} of Mns was calculated to be $1.00x10^5$ +/-926 cal mol^{-1}, and is close to the absolute enthalpy ΔH_{cal}. Agreement of these estimates is indicative of a two-state reaction (van Nuland *et al.*, 1998).

3.4.8.3 Thermal Stability

The melting temperature of Mns presented here approaches that observed for the proteins of thermophilic archae (**Table 3.1**). This is particularly surprising because *P. luminescens*, which also produce Mns, grow optimally at 28°C and are considered to be facultative psychrophiles.

There is no available benchmark of thermal stability of *Photorhabdus* proteins; Manukhov *et al.* (1999) have focused on the stability of the luciferase proteins, which lose enzymatic activity above 37°C and fail to recover from extended (1h) heat shock treatment at 47°C. Similarly, Bowen *et al* (2003) have exposed the *P. luminescens* protease PrtA to different temperatures. The protein undergoes calcium-dependent cleavage and autocleaves at 42°C and 60°C, but shows remarkable stability at 37°C. No melting temperature data are available for PrtA. For *P. asymbiotica*, there is a necessity for proteins to remain structurally stable at mammalian physiological temperatures. Stability above these temperatures must be attributed to the sequence and structural features of the protein.

An analysis by Kumar *et al.*, (2000) of the structural stability of proteins from thermophilic organisms with homologs from mesophiles shows differences in the distribution of amino acid residues. Cysteine is considered to be a thermolabile amino acid due to its tendency to undergo oxidation at high temperatures, and appears to be avoided in proteins from thermophilic archea and bacteria. Serine residues are also less frequent, while arginine and tyrosine are more abundant and are expected to contribute to thermal stability by making both long-range and short-range interactions due to their long side chains. In Mns, cysteine accounts for just 0.7% of all residues, consistent with expectations for thermostable proteins. However, serine is the most abundant of all residues (13.2%).

Differences also exist between the fraction of residues existing in helical conformations in thermophilic and mesophilic proteins; the thermostable proteins are richer in α-helices than mesophilic proteins (Kumar *et al*, 2000). Furthermore, there are differences in the residue composition of these helices, with cysteine and histidine, helix-disfavouring residues, significantly less populous in the helices of thermophilic proteins. Proline, which may cause kinks in the helix, is the most avoided residue in the middle of α-helices and is also less frequent in these proteins. By contrast, arginine is a helix-favouring residue and is more frequent in the helices of thermophilic proteins. It would be of interest to identify which residues form the predicted 7-8 α-helices of Mns.

Thermal stability therefore appears to result from factors that enhance secondary structure propensity; thermophilic proteins tend to have more residues that favour α-helices and have fewer residues that destabilise these conformations (Querol *et al.*, 1996). The stability of Mns likely results from its α-helical richness, and the content of these helices. Kumar *et al.*, (2000) found no correlation between thermal stability and changes to overall protein size by the deletion or shortening of loops or by oligomerisation. However, in addition to sequence factors, structural features such as packing, polar surface area and hydrogen bonds were found to enhance protein stability. Salt bridges also contribute, probably by opposing the disorder forced on the residues by greater mobility at high temperatures. Without high-resolution structural analysis we are unable to explore these contributions in Mns.

Finally, it is noteworthy that the living temperature of an organism is not a good descriptor of thermal stability, and melting temperatures are more appropriate for evaluating the stability of a protein (Kumar *et al.*, 2000). It follows that one should take caution in making a prediction about the thermal stability of a protein by considering only the temperature of its surroundings; it is unlikely that *Photorhabdus* will encounter temperatures above 40°C, but the melting temperature of Mns above 60°C should not be unexpected.

3.4.9 Congo Red Absorbance and Amyloid Fibril Formation

Our observation that massive overexpression of recombinant Mns in *E. coli* cells resulted in cellular elongation and lack of septation with the formation of long fibrous material inside the cells suggested that the protein may form amyloid-like aggregates. We also observed that nematodes aggregated by *P. asymbiotica* supernatant stained with the dye Congo red, and Western analysis showed that Mns was present in these aggregates.

To test the hypothesis that Mns may form amyloid fibrils, purified recombinant protein was incubated with Congo red. Congo red is used to identify amyloid fibrils in samples of diseased tissue, such as in Alzheimer's and prion diseases. Binding of the dye to amyloid fibrils is associated with a spectral change and a maximum difference in absorbance at the wavelength 541nm. Following the method of Chapman *et al.*, (2002) we were unable to detect this spectral change with Mns, while a large difference was seen with prion protein prepared to the same concentration.

This test may have suffered from a lack of necessary cofactor(s). It has been reported elsewhere that proteins that form amyloid fibril deposits are often associated with other nonfribrillar components, such as proteoglycans or lipoproteins (Sipe & Cohen, 2000). A thorough screen may be required to identify such a component. Alternatively, different solution conditions such as pH and ionic strength should be explored to find an appropriate condition to promote fibril formation (Nilsson, 2004). If such a condition could be found it would be of particular interest to study the circular dichroism of the protein, which would be expected to reveal a secondary structure rich in β-sheets. Lastly, it is also noteworthy that Congo red does not bind all amyloid deposits by the same mechanism, and a particular amino acid composition

or molecular arrangement of packing may prevent detection by the dye (Nilsson, 2004).

3.4.10 Conclusions

The structure of Mns remains to be solved. Circular dichroism spectroscopy has provided convincing evidence that Mns is a helix-rich protein, but also contains β-sheets and is an α+β type protein. Secondary structural prediction programmes isolate helices toward the N-terminus and suggest, contrary to CD data, that the majority of secondary structural elements are β-sheets. Prior to aggregation, Mns gave excellent chemical shift dispersion in initial NMR spectra, consistent with the protein possessing substantial β-secondary structure.

Dynamic light scattering and analytical gel filtration suggest that Mns exists as a dimer. Through DLS and NMR data collection, we have also shown that the protein has a high propensity to aggregate, a feature that would be particularly consistent with a β-sheet enriched structure. Potential aggregation sites have been identified in the protein, and these correspond to predicted β-sheet regions. Unfortunately, CD is unable to identify which residues contribute to specific secondary structural elements, and the propensity of Mns to aggregate has impeded our efforts to investigate these features by NMR. Further analyses are required to examine the possibility that Mns undergoes conformational changes, such as from a predominantly helical structure to a β-rich structure, and the conditions in which this change occurs. Finally, our differential scanning calorimetry data show that Mns has an exceptionally high melting temperature. Solving the structure of Mns will provide an explanation for this stability.

CHAPTER 4

General Discussion and Future Perspectives

4.1 GENERAL DISCUSSION

Our observation that Mns has a tendency to aggregate hampered structural analysis but may have implications to the proteins function. The role of Mns appears to relate to its 'stickiness', or its ability to modify attachment. Stickiness has been described elsewhere in relation to cellular attachment and the formation of biofilms. For instance, in *S. marcescens*, a pathogen of both humans and insects, a number of phenotypic 'sticky colony' variants have been identified. Two of these variants, SSV and SRV, attached no better than parental wild-type cells to microtitre plates, but were able to form two- to threefold more biofilm (Koh *et al.*, 2007). Interestingly, for a third variant (SRUV), significantly less biofilm formed on microtitre plates, while copious amounts of biofilm formed in a flow cell system, suggesting that physiological conditions differ between the two systems that favour or inhibit biofilm formation (Koh *et al.*, 2007). Clearly, inside an insect a variety of different environments may be encountered, and these observations validate our data for Mns, in which Mns-knockout cells were shown to attach better than wild-type cells to gold surfaces, but formed less biofilm growth on polypropylene plates. While it should be pointed out that we have indirect evidence for the attachment of Mns to both gold and polypropylene, revealed by SPR of anti-Mns antibodies and Western blot analysis, respectively, the subsequent attachment of cells to the Mns-modified surfaces may be under the control of a number of factors. Furthermore, the techniques used to assess biofilms have some bias, e.g. SPR measures attachment of all materials while growth is a microtitre plate is underestimated because the crystal violet used to detect biofilm does not stain extracellular polymeric substances (Koh *et al.*, 2007).

To further characterise the protein, assays of colonisation of biological surfaces should be pursued. Although Mns was not found to be toxic, and no observable differences were seen in *G. mellonella* infected with wild-type and Mns knockout *P. luminescens*, there may be important differences in the colonisation of these insects. For instance, the presence of Mns may enhance attachment to the fat body or to epithelial cells of the insect midgut, or allow the bacteria to survive in a particular niche. The high conservation of Mns suggests that any interaction with insects is not

species specific, and our use of two lepidopteran model insects is unlikely to have restricted our ability to find a function in pathogenesis.

The 'stickiness' of Mns we have observed, particularly with native protein in *P. asymbiotica*, but also through aggregation of recombinant Mns, may give clues to the role it plays in *Photorhabdus* ecology, as well as give insights to its structure. nFGF-1 is a 15kDa protein that adopts a native β-barrel conformation and is able to form amyloid-like fibrils. In high concentrations of trifluoroethanol, the protein is induced to form helical conformations (Srisailam *et al.*, 2003). Importantly, an intermediate stage is recognised, in which the β-barrel conformation becomes disorganised, resulting in the formation of extended β-sheets. This state, which has a non-polar surface exposed to solvent, is described as 'sticky' and allows maximal aggregation of the protein into fibril formations. Furthermore, conditions that promote helical conformations are associated with decreased aggregation in nFGF and other proteins (Srisailam *et al.*, 2003, Chiti *et al.*, 2003). Although the aggregation of Mns has been shown in the present study not to form amyloid fibrils, it appears to form a fibrous material, and our different estimates of protein structure deriving from NMR and CD suggest a possibility that the protein undergoes secondary structural changes from an α-helical form to a β-rich form. In the case of massive overexpression in *E. coli*, the fibril-like material may be a result of molecular crowding causing aggregation (Ellis, 2006). The details of aggregation differ between proteins, such as the residues important in forming specific interactions, but more importantly there appears to be set of common principles, such as residue hydrophobicity, charge, and propensity to change from α-helical to β-sheet structures (Chiti *et al.*, 2003). Perturbation of these parameters can result in increased or decreased rates of aggregation. For example, in α-synuclein the mutation A53T results in accelerated aggregation and is associated with pathogenicity (Conway *et al.*, 2000), while in Mns, a number of mutations have been identified (G. Yang, unpublished data), such as I93N, that do no not result in the formation of cell chains in *E. coli*, presumably because of reduced aggregation propensity. Whether Mns does indeed undergo any structural changes has yet to be determined, but the aggregation we have observed is likely to occur in the native conformation and is central to its role in modifying cell attachment.

The potential ability of Mns to bind to sugars is particularly intriguing, and is consistent with our observation that Mns in colony biofilms was bound to an extracellular material that is likely to be *Photorhabdus* exopolysaccharide. Together with our SPR and cell attachment data, it could be interpreted that Mns binds the extracellular material that surrounds cells and modifies its adhesive properties. This could have implications in both the symbiotic and pathogenic stages of the *Photorhabdus* lifecycle. Furthermore, binding to sugars on the nematode cuticle could explain the interaction of Mns with *C. elegans*, as has been proposed for *M. nematophilum* infections of nematodes.

Since Mns is expressed at high levels in both the exponential phase and stationary phase, we cannot categorise the protein on the basis used by Joyce & Clarke (2003) and assign a role for the protein in symbiosis or pathogenesis. Our finding that Mns is not toxic to insects when administered orally or by injection suggests that the protein does not function in the same was as its closest known homolog, the binary crystal toxins. However, the caveat mentioned earlier may be significant; we were only able to purify recombinant Mns and the toxicity of the native protein has not been examined. Indeed, it may be significant that the native protein was not amenable to purification, and perhaps suggests some modification that has not occurred in Mns expressed heterologously in *E. coli*.

4.2 FUTURE PERSPECTIVES

To further understand the role of Mns in insect infection, future studies with Mns can concentrate firstly on characterising the expression profile of the protein. We isolated midgut, fat body and haemolymph from *M. Sexta* after 24h to ensure that these tissues could be efficiently separated, but were unable to detect Mns. Repeating this study with later time points, but still prior to extensive tissue damage, may show that the protein is differentially expressed in insect tissues. It may also be necessary to try an extensive analysis of pathogenicity of wild-type and Mns- knockout bacteria to search for a role in pathogenicity. For instance, injecting several dilutions of cells, or strains with other pathogenicity factors knocked out, may allow subtle differences in virulence to be identified.

Further studies are underway to explore the possibility that Mns is a sugar binding protein. In addition to dot-blot analyses with exopolysaccharides, mass spectrometry is being used to test for differences in molecular weight from the hypothetical weight. If successful, NMR samples may be prepared with sugars introduced to look for differences in spectra and to see if the aggregation of the protein is affected. Purifying native Mns from the supernatant of *P. asymbiotica* has thus far been unsuccessful, but would help to answer a number of question raised in the present study. Finally, the study of interactions between *Photorhabdus* and *C. elegans* in this study has also yielded an observation of nematode paralysis. This occurred with cultures incubated for extended periods and was distinct from the nematode aggregation phenotype. Interestingly, *C. elegans* paralysis has been observed previously, resulting from a diffusible toxin produced by *P. aeruginosa* (Darby *et al.*, 1999). The paralysis factor of *P. aeruginosa* has not been identified, but appears also to cause bronchoconstriction in a mammalian infection, indicating that the toxicity factor may be of significance in human infections. Bacterially induced nematode hypercontractile paralysis has also been reported for *Burkholderia pseudomallei* (O'Quinn *et al.*, 2001) and enteropathogenic *E. coli* (Anyanful *et al.*, 2005). The growing number of human infections demands that these paralysis factors and their targets be identified for *P. asymbiotica*, as well as for the more-established human pathogenic bacteria.

CHAPTER 5

CYP6G1-Mediated Insecticide Resistance in *Drosophila melanogaster*

5.1 INTRODUCTION

Insecticides are central to the control of agriculturally important arthropod pests and insects of medical significance, particularly those that are vectors of diseases. However, resistance to insecticides can occur in any species and has become a major constraint to the control of these insects because it results in less effective toxins, driving increases in application frequency and dosage, or the design of new toxins in severe cases when the utility of entire classes of insecticide is lost.

Furthermore, because resistance is an inherited ability to survive doses of the toxicant that kill the majority of individuals in a normal field population of the same species, the genetic factors responsible have become widespread and represent an important example of anthropogenic natural selection.

5.1.1 Types of Insecticide and Resistance

The use of insecticides was revolutionised following the Second World War with the introduction of small molecule synthetic insecticides. The first of these, DDT, became the most useful insecticide developed to date. Later came other organochlorines such as dieldrin, carbamates, organophosphorous compounds and pyrethroids. Safety concerns and widespread resistance has since depleted the arsenal of available insecticides (Ware, 2004).

Resistance can be engendered through behavioural actions (avoidance of exposure) or physiological means (survival of exposure). Physiological resistance includes reduced penetration of the toxin, increased excretion and increased metabolic detoxication. These mechanisms are achieved by (i) constitutive overproduction of a gene product; (ii) constitutive underproduction or (iii) an inducible change in gene regulation. A further class exists of insects whose resistance to a xenobiotic is acquired physiologically through an alteration of the target site.

5.1.1.1 Target Site Insensitivity

The major target sites of commonly used pesticides are the Rdl (resistant to dieldrin) subunit of the γ-aminobutyric acid receptor (GABAR), the target of cyclodiene and fipronils; the *Ace*-encoded acetylcholinesterase (AChE), target of carbamate and organophosphorous; and the voltage-gated sodium channel protein PARA.

The voltage-gated sodium channel is required for transmission of nerve impulses and is the target of at least nine distinct categories of neurotoxins from plant and animal allomonal systems, as well as four classes of industrial insecticides; DDT, N-alkylamides, dihydropyrazoles and pyrethroids (Bloomquist *et al.*, 1996). The knockdown resistance (KDR) phenotype, in which the insect nervous system may be 50-100-fold less sensitive to the insecticide, has been associated with the *para*-like sodium channel gene. In houseflies, two amino acid substitutions in domain II of the α subunit of the sodium channel have been identified in KDR strains (Williamson *et al.*, 1996). These mutations are hypothesised to alter the binding site of the insecticidal toxin and, despite the complexity of the receptors, are extremely conserved among a wide range of insects selected by different insecticides. Similarly, a single residue is replaced in Rdl (ffrench-Constant *et al.*, 1993).

Resistance to organophosphates is associated with changes in AChE, an enzyme in cholinergic synapses required to terminate nerve impulses. Organophosphates are substrates of AChE and their hydrolysis blocks synaptic transmission. Pests can acquire resistance through modification of the enzyme; three or four mutations in the gene encoding AChE provide high level and wide-spectrum resistance as the enzyme is much less inhibited than that from the susceptible strain (Menozzi *et al.*, 2004).

Numerous mutations to the target site of the insecticide may potentially confer resistance to the organism, but studies of insect populations frequently find very few alleles. This is likely due to the requirement for resistant targets to maintain their wild-type function, and the physiological cost associated with the mutations; the alleles associated with the lowest fitness cost in the absence of the insecticide are selected (Shi *et al.*, 2004). Modifier genes, at the same or different loci, that restore the initial activity of the mutated protein and subsequently restore the fitness of the

insect, may reduce the fitness cost. A typical cost of resistance is a change in an enzyme or receptor with some disruption to its optimal functioning. Another cost might be the over-expression of a receptor or enzyme, diverting resources away from other fitness-enhancing activities (Coustau *et al.*, 2000).

5.1.1.2 Metabolic Detoxification

The transcriptional overexpression of metabolic enzymes allows many insect populations to detoxify the large variety and quantity of toxins to which they have become exposed since the widespread introduction of synthetic organic agrochemicals. The enzymatic breakdown of insecticides to more hydrophilic and less toxic molecules is possibly the most commonly occurring mechanism of resistance, and has the potential to confer cross-resistance to toxins independent of their target site. The increased metabolic activity is frequently due to upregulation of esterases, glutathione transferases and, primarily, cytochromes P450.

5.1.2 Cytochromes P450

The cytochromes P450 (P450s) form a superfamily of heme-containing monooxygenases. These enzymes are extremely important because of their role in the metabolism of endogenous compounds such as steroids, bile and fatty acids. Many also metabolise insecticides and a wide range of other exogenous toxins (Feyereisen, 1999).

P450s reductively cleave molecular oxygen, producing an organic product and a molecule of water;

$$RH + O_2 + 2e^- + 2H^+ \rightarrow R\text{-}OH + H_2O$$

This monooxygenase-mediated oxidation requires P450, NADPH-cytochrome P450 reductase (P450 reductase), NADPH, phospholipid and in some instances cytochrome b_5 (Lu *et al.*, 1974). Many P450s are not specific and are capable of hydroxylating a variety of diverse and unrelated compounds, and as little as a single amino acid substitution can considerably alter the substrate specificity of the enzyme. Furthermore, the number of products from a given substrate ranges from a single to multiple metabolites.

P450s are ubiquitous in nature and are found in almost all aerobic organisms. The gene superfamily now comprises over 850 families, many divided further into subfamilies (Nelson, 2007). All gene members are designated with a *CYP* prefix, followed by a numeral, a letter and a numeral for the family, subfamily and individual gene (Nelson *et al.*, 1996). Although the best-studied P450 systems are mammalian, over 100 different P450 genes have now been identified in insects and are involved in pathways central to insect growth, development and reproduction. An association between insect P450s and resistance to insecticides was forged in 1960, when Eldefrawi *et al.* demonstrated the abolition of carbomyl resistance by inhibiting P450 activity.

Cytochromes P450 are now seen as a major class of enzymes involved in the detoxification of xenobiotics. Insect genomes carry about 100 P450 genes, each encoding a different enzyme (Feyereisen, 1999). The recent publication of the *Apis mellifera* genome shows that the honeybee has just 46, with particular shortfalls in P450s implicated in insecticide metabolism. The finding that honeybees are relatively depauperate in P450 enzymes may offer an explanation for their susceptibility to pesticides (Honeybee Genome Sequencing Consortium, 2006).

5.1.3 Constitutive Expression of Cytochromes P450

The induction of metabolic enzyme expression by the substrates they act upon has been characterised in many biological systems, and allows many organisms to enervate the toxic compounds they encounter, e.g. in the lepidopteran *Papilio polyxenes*, the P450 genes *Cyp6b1* and *Cyp6b3* are overexpressed upon exposure to xanthotoxin, a furanocoumarin found in their host plants, and are responsible for its detoxification (Petersen *et al.*, 2001).

By contrast, there appears to be little induction response to insecticide exposure. Willoughby *et al.* (2006) found only one *D. melanogaster* P450 gene was induced when testing with six different insecticides, and the gene (*Cyp12d1*) had a much higher level of induction with the barbituate drug phenobarbital and the plant compound caffeine. To achieve P450-mediated resistance, the predominant response appears to be a high level of constitutive expression.

The identification of the particular P450 gene or genes linked with resistance is often difficult. However, in cases where a correlation has been found, *Cyp* genes with elevated activity are mostly, but not exclusively, members of the *Cyp6* family. Further to the examples above, deltamethrin resistance in the mosquito *Culex pipiens pallens* appears to involve up-regulation of *Cyp4* family members (Shen *et al.*, 2003), while permethrin resistance in *Culex pipiens quinquefasciatus* has been associated with elevated transcript levels of *Cyp6f1* (Kasai *et al.*, 2000); the over-production of CYP6D1 and CYP6Z1 is observed in pyrethroid-resistant isolates of *Musca domestica* and *Anopheles gambiae,* respectively (Kasai and Scott, 2000, Nikou *et al.*, 2003).

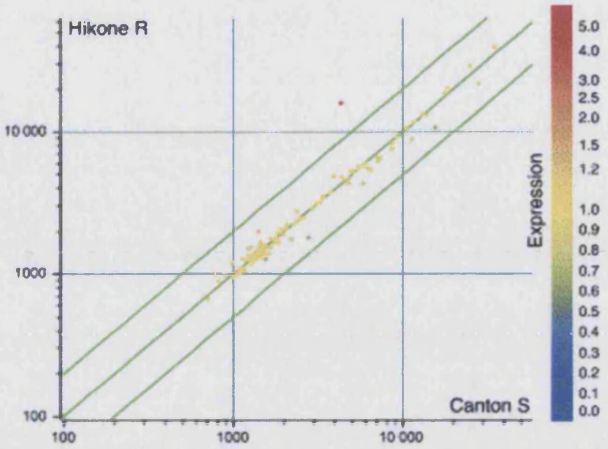
5.1.4 *Drosophila* P450s

The fruitfly *Drosophila melanogaster* is increasingly used as a model organism for the study of insecticides as both a problem and as a biologically interesting occurrence. This is due to the genetic resources available for the species, and discoveries that field populations have evolved resistance to toxic chemicals released into the environment by humans.

The entire *D. melanogaster* genome has been sequenced, and contains 90 cytochrome P450 genes (Tijet *et al.*, 2001). Several of these genes have been implicated in conferring resistance to insecticides, including *Cyp6a8* and *Cyp6a9* (Maitra *et al.*, 1996 and 2002). In the case of *Cyp6a2*, resistance is mediated not from a genetic change that results in altered gene expression, but to a change in enzyme function. Point mutations have altered the substrate profile of the enzyme, allowing DDT to be metabolised and resistance to be established in the $RDDT^R$ strain of flies (Amichot *et al.*, 2004).

A microarray analysis of laboratory selected strains with resistance to DDT shows upregulation of a range of P450s, while in recent field isolates resistance is conferred only by gene, *Cyp6g1* (Le Goff *et al.*, 2003). Transgenic flies overtranscribing a copy of *Cyp6g1* showed that the gene is both necessary and sufficient for resistance (**Fig. 5.1**) (Daborn *et al.*, 2002b). Furthermore, the *Cyp6g1*-transgenic flies showed cross-resistance to malathione, neonicotinoids, organophosphorous compounds and lufenuron.

(A)



(B)

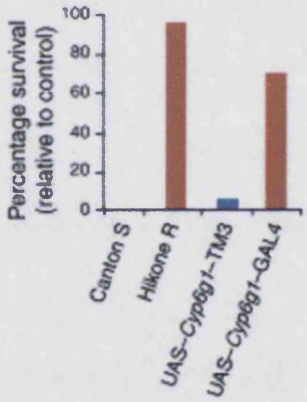


Fig. 5.1 Implication of *Cyp6g1* in DDT resistance. **(A)** Transcriptional analysis using a DNA microarray of all 90 *D. melanogaster* P450 genes using RNA isolated from resistant and susceptible strains. The array shows the relative hybridisation intensities for each of the P450s in the resistant strain (Hikone-R) against the susceptible strain (Canton-S). The only P450 gene over-transcribed in the field-isolated resistant strain is *Cyp6g1* (red spot). **(B)** The relative survival of flies of different genotypes treated with DDT. Resistant (Hikone-R) flies, or flies carrying both UAS–*Cyp6g1* and GAL4 (over-transcribing *Cyp6g1*), can survive insecticide exposure (red bars); the standard susceptible strain Canton-S, and flies lacking GAL4 (blue or no bars), can not, providing evidence that over-transcription of *Cyp6g1* can confer DDT resistance. (All data from ffrench-Constant *et al.*, 2004)

The use of DDT to control insects has been banned, and yet field isolated samples maintain their resistant phenotype. Furthermore, *Cyp6g1*-mediated resistance to DDT can persist in laboratory strains in the absence of insecticide selection. In these flies, resistance correlates with the presence of an *Accord* transposable element inserted into the 5' region of the gene (Daborn *et al.*, 2002b). This resistant allele appears to have formed following a single transposition event into one susceptible haplotype, and has subsequently spread globally. That a single change in a single gene in the field has been selected likely reflects the infrequency of a mutation causing a viable resistance phenotype. It is further consistent with a hypothesis in which a P450 conferring broad cross resistance has a selective advantage over other P450 up-regulation genotypes because it may be selected by a wider range of compounds (Le Goff *et al.*, 2003).

Lastly, it is noteworthy that a *Cyp6g1* allele with a different transposable element has been identified in a geographically restricted population of *Drosophila simulans* (Schlenke and Begun, 2004). In a remarkable example of parallel evolution, the insertion of the *Doc* element confers a degree of resistance to DDT, just as the *Accord* element does to *D. melanogaster*.

5.1.5 CYP6G1 and Hormone Metabolism

In addition to the cross-resistance that over-expression of *Cyp6g1* confers to *D. melanogaster*, Hikone-R flies have enhanced adult fecundity and egg and larval viability. The flies also show significantly faster rates of larval and pupal development (McCart *et al.*, 2005). These findings suggest that *Cyp6g1* overexpression is selected not just because it provides resistance to several insecticides, but because it confers a fitness advantage to flies. The data also counter the current theory of resistance carrying a fitness cost, and invite speculation that *Cyp6g1* may be involved in insect development through the metabolism of hormones. There is little distinction between insect P450s involved in the detoxification of xenobiotics and those involved in physiological pathways. Indeed, CYP6A1 from the house fly *Musca domestica* has an unusual active site topology that allows for the metabolism of both pesticides and the terpenoid precursors of the juvenile hormones (JH), with little substrate or product specificity (Andersen *et al.*, 1997).

Showing a higher level of stereoselectivity, CYP15A1 from the cockroach *Diploptera punctata* has been identified by Helvig *et al.* (2004) as a P450 that catalyses the epoxidation of methyl farnesoate. This is the final step in the production of the sesquiterpenoid juvenile hormone III. The enzyme is expressed in the specialized endocrine gland the corpora allata, and titres of its product are central to the control of the insect's maturation, including maintenance of the juvenile state and development of the reproductive organs.

Other developmental events, chiefly the periodic moulting of the exoskeleton, are controlled by the moulting hormones, the ecdysteroids. The final four sequential steps in ecdysteroidogenesis are mediated by cytochromes P450. The Halloween P450s Phantom (CYP307A1), Disembodied (CYP302A1) and Shadow (CYP315A1) are expressed in the prothoracic glands, and complete the pathway that produces ecdysone (Rewitz *et al.*, 2006). The hydroxylation to 20-hydoxyecdysone (20E), the principle insect moulting hormone, is achieved in the fat body, midgut and Malpighian tubules by Shade (CYP314A1). Like other Halloween mutants, *shade* mutants do not produce a differentiated cuticle and late embryonic morphogenesis is severely disrupted (Petryk *et al.*, 2003).

Potential crossover between xenobiotic detoxification and hormone metabolism exists where insecticides have been synthesised that mimic insect hormones, such as the ecdysteroid agonist tebufenozine and methoprene which mimics JH, and where plants produce phytoecdysteroids as agents to prevent herbivory (Savchenko *et al.*, 2000). Whether CYP6G1 is capable of metabolising insect hormones has not been tested but deserves attention. The current study approaches this question, but focuses on the ability of the protein to confer a broad range resistance to insecticides. Elucidation of how CYP6G1 and other P450s accommodate their substrates may be best achieved through studies of the protein's molecular structure.

5.1.6 Objectives

Determining the molecular basis of resistance is important to studies of population genetics and evolution, and to the design of effective countermeasures. In the present study, the molecular structure of CYP6G1 is examined by homology modelling. The structures of CYP12D1 and CYP6A2 are also studied with particular reference to their substrate profiles. Lastly, the expression of CYP6G1 in *D. melanogaster* is probed, and the potential role of the enzyme outside of insecticide metabolism is discussed.

5.2 MATERIALS AND METHODS

5.2.1 Cloning *Cyp6g1* into Parallel Vectors

Oligonucleotide primers were designed for polymerase chain reaction (PCR) amplification of *Cyp6g1* from its initial cloning vector, pUAST. The *Cyp6g1*-pUAST clone was kindly provided by S. Boundy. The primers were designed for introduction of the entire gene into the maltose binding protein (MBP) and glutathione-S-transferase (GST) tag parallel vectors (Sheffield *et al.*, 1999). In each vector, a tobacco etch virus (TEV) protease cleavage site was encoded between the affinity tag sequence and the multiple cloning site into which the gene was inserted.

The *Cyp6g1* gene was cloned using a *Bam*HI restriction site introduced at the 5' end and a *Not*I site incorporated at the 3' end of the gene. The cloning procedure was the same as that described below for cloning into pBluescript, with the exception that *Bam*HI and *Not*I were used in the sequential digestions, blue-white selection was not used, and plasmids were transformed into EC100 cells.

5.2.2 Cloning *Cyp6g1* into pBluescript and pET15b

The PCR product for cloning into pET15b was designed to amplify the entire coding region, with the exception of an identified N-terminal transmembrane domain. The first primer set (Cyp6g1-T1) omitted the first 63 bases encoding this domain, and the second set (Cyp6g1-M1) omitted the first 63 bases and introduced a short sequence encoding hydrophilic residues to modify the N-terminus.

The primers were also designed to incorporate restriction endonuclease cleavage sites: a *Kpn*I restriction site at the 5' end and a *BamH*I site at the 3' end of the gene. Inside the *Kpn*I site an *Nde*I site was fitted, to allow insertion of the gene into pBluescript II KS+ for blue-white selection, then excision and ligation into the *Nde*I site of pET15b for expression.

For each reaction, 10µl each of forward and reverse primer (10pmol/µl) was mixed with 10µl 2.5mM dNTP mix, 1µl plasmid DNA, 1µl Taq polymerase and 10µl Taq Buffer (New England Biolabs). Two concentrations of $MgCl_2$ were used to optimise PCR output: 4.8µl or 6.4µl of 25mM stock were added, and the reactions were made up to 100µl with sterile distilled water. The thermal cycle incorporated a 95°C denaturation step (2min) followed by 36 cycles of denaturation (95°C, 30s), annealing (50°C, 30s) and extension (72°C, 3min). A 72°C extension (10min) completed the programme.

Excess primers and nucleotides were cleaned from the PCR products with a Montage cleanup kit (Millipore), following the manufacturers instructions. For restriction digestion, 40µl of purified *Cyp6g1* PCR product was digested with 2.5µl *BamH*I (NEB), with 5µl *BamH*I buffer and 2.5µl sterile distilled water. The reaction was incubated at 37°C for 2h and impurities were removed with the Montage kit. The eluted DNA was digested in the same way with *Kpn*I. The pBluescript II KS+ (Stratagene) vector was also sequentially digested.

The *Cyp6g1* insert and vector DNA were separated on a 0.7% agarose gel at 100V to confirm the size of the products and approximate concentrations before ligating in a ratio of 3:1 insert:vector. 1µl ligase, 5µl rapid-ligation buffer (Promega) and 4µl of insert/vector were mixed and incubated at room temperature for 10min. 5µl was added to 50µl EC100 competent cells (Epicentre) for electroporation at 200Ω, 25µF capacitance and 2.5V. Cells were recovered in 500µl LB, incubated for 1h at 37°C and plated on LB amp agar supplemented with isopropyl thiogalactoside (IPTG) to a final concentration of 0.1mM and Xgal to a final concentration of 40µg/ml.

Cells transformed with the pBluescript vector containing an inserted DNA fragment were identified by a lack of blue colouration, and colonies were picked to inoculate 3ml LB amp broth. Cloning of *Cyp6g1* was confirmed by plasmid preparations (Miniprep kit, Qiagen), which were digested with *Kpn*I and *BamH*I and run on an agarose gel alongside the *Cyp6g1* insert and digested vector. The clones were sequenced from the vector using the T3 and T7 promoter sites, and internally (primers provided by S. Boundy).

Plasmid preparations of cloned and sequenced *Cyp6g1* were finally digested with *BamH*I and *Nde*I and ligated as described above into the pET15b vector (Novagen). For expression, the ligation was transformed into DE3 Rosetta competent cells (Novagen) following the manufacturer's instructions, and transformants were selected on LB amp, chloramphenicol (cam).

5.2.3 Heterologous Expression of CYP6G1

For expression of CYP6G1-M1, -T1 and CYP6G1-MBP, 1ml of overnight cultures containing the plasmid constructs grown in selective media were pelleted by 2min, 16,000x g centrifugation and resuspended in 100ml fresh selective broth. The cultures were incubated at 37°C, with shaking at 180rpm to an OD_{600} of 1.1, induced with IPTG (0.1mM final concentration), supplemented with δ-aminolevulinic acid (ALA) to a final concentration of 0.5mM and returned to incubation at 37°C. The optical density was measured at hourly intervals and a sample taken to monitor expression (as described in Chapter 2).

Subsequent expression trials included an overnight induction up to 24h, growth in either LB or Terrific Broth (Sambrook *et al.* 1989), incubation at 18°C, 28 °C or 37°C, and in the case of CYP6G1-MBP, transformation of the expression vector into Rosetta and BL21* cells (Novagen). Protein solubility tests were performed with TCA (as described in Chapter 2, using PBS as a resuspension buffer for sonication). To improve solubility, different lysis buffers were tested using a range of pH conditions (sodium acetate pH 5.5, Tris pH 7.0, Tris 7.4 and Tris 8.0), as well as 0.1mM DTT, 0.01% SDS and 0.1% Triton X additives and 10% glycerol.

5.2.4 Purification and Cleavage of CYP6G1-MBP

CYP6G1 expressed as a fusion protein from the MBP-parallel vector was produced in a 100ml culture grown at 18°C and induced as above. The culture was pelleted at 15,000x g, 10min, 4°C. The pellet was resuspended in 25ml Wash buffer (20mM Tris HCl pH 7.0, 0.2M NaCl) and sonicated. For purification using the MBP tag, a 10ml column was washed and packed with amylose resin (GE Healthcare). The resin was washed with 2x column volumes of water, then 2x column volumes Wash buffer. 50ml of soluble fraction, obtained after centrifuging sonicated cells, were passed through the column at an approximate flow rate of 1ml/min. The flow through was

collected and 2x column volumes Wash buffer were passed though the column before Elution buffer (20mM Tris HCl, pH 7.0, 0.2M NaCl, 10mM maltose) was introduced at 1ml/min. 30x 1ml fractions were collected. The optical density of the eluted fractions was measured at 280nm and plotted to identify an elution peak.

Batch purification of CYP6G1-MBP was also performed. The pellet of 1 litre of induced culture was resuspended in 25ml of a no-salt Lysis buffer (20mM Tris HCl, pH 7.0, 0.5mM EDTA and 0.1mM DTT). Following sonication as described, the supernatant was added to amylose resin equilibrated with Lysis buffer, in a 50ml conical tube. (To equilibrate the beads, approx 10ml of amylose resin, with a bed volume of 3ml was centrifuged at 800x g, 10 min. The supernatant was removed by aspiration and the beads were washed in 3x bed volume of Lysis buffer. The resin was pelleted as before and washed a further two times). The sample was placed on a rocker at room temperature for 1h, and pelleted (800x g, 10 min). The supernatant was removed and the beads were washed three times with 3x bed volume TEV cleavage buffer (50mM Tris HCl, 0.5mM EDTA). Finally, the beads were resuspended in TEV cleavage buffer and a 10µl sample was added to 3µl 3x SDS loading buffer for SDS-PAGE.

5.2.5 Fusion Protein Cleavage

Partially purified CYP6G1-MBP protein was cleaved with the recombinant TEV protease (Invitrogen). 7.5µl of 20x rTEV Buffer (1M Tris-HCl (pH 8.0), 10mM EDTA), 1.5µl 0.1M DTT and 0.1µl (1 unit) TEV protease were added to 20µg of fusion protein, and made up to 150µl water in a microcentrifuge tube. All reactions were performed at this volume, and left overnight at room temperature. SDS-PAGE of 10µl of cleavage reaction boiled for 5min with 3µl of 3x SDS-loading buffer revealed the extent of cleavage. The cleaved product was separated from MBP by passing through an amylose column.

For cleavage of batch-purified protein, the addition of DTT to a final concentration of 0.1mM to the fusion protein bound to amylose beads was followed by addition of TEV protease (30µl to a 50ml tube with a bed volume of 3ml). The reaction was allowed to proceed overnight on a rocker at RT. The sample was then pelleted (800x g, 10 min) to separate the target protein from the resin-bound MBP. The beads were also washed with TEV buffer to further remove the cleaved protein, and the samples were analysed with SDS-PAGE.

5.2.6 Purification and Refolding of CYP6G1 for the Purification of Anti-CYP6G1 Antibodies

CYP6G1 was purified using the N-terminal hexahistidine tag introduced by pET15b for the purpose of purifying anti-CYP6G1 antibodies away from a polyclonal serum. The presence of the haem group in the active site was therefore not important and the protein could be denatured and refolded to improve solubility and allow purification.

Cultures were grown in selective LB and were incubated at 18°C overnight after induction with IPTG (1mM). The pellet of 600ml was resuspended in 5ml Solubilisation buffer (50mM Tris pH 8.0, 300mM NaCl, 6M guanidine HCl) and sonicated. The lysed cells were pelleted at 16,000x g, 30min, 4°C and diluted to 20ml with Solubilisation buffer.

A Sartobind IDA 75 metal chelate membrane adsorber (Sartorius) was recharged following the manufacturer's instructions and loaded with the soluble fraction after filtering with a 0.45µm filter (Millipore). The flow through was collected and the membrane washed with 10ml Solubilisation buffer before eluting with 5ml solubilisation buffer containing 250mM imidazole. The column was washed with water and re-equilibrated before re-loading the flow through and repeating the elution process.

The eluted products were pooled and added dropwise over 2-4h into 250ml stirred Refolding buffer (0.5M L-arginine, 0.6mM oxidised glutathione, adjusted to pH 8.0 with HCl). The refolded protein was left at room temperature overnight before concentrating in a 10,000MWCO centrifuge concentrator (Millipore). The concentrated protein was analysed by SDS-PAGE alongside uninduced and induced whole culture samples (pellet of 500µl culture resuspended in 50µl 1xSDS gel loading buffer, loaded as $1/OD_{600}$ x 4µl).

5.2.6.1 Preparation of Polyclonal Antibody Serum for Purification

To remove serum proteins from polyclonal antibodies raised in rabbits against a CYP6G1-maltose binding protein (MBP) fusion protein (provided by S. Boundy), a 1ml protein A column (GE Healthcare) was prepared. The column was washed by syringe injection at approximately 1ml/min with water (5 column volumes), elution buffer (100mM sodium citrate, pH 4.0, 5 column volumes) then binding buffer (20mM sodium phosphate, pH 6.7, 10 column volumes). 2ml of polyclonal antibody serum was diluted with 8ml binding buffer and loaded by syringe onto the column. The column was washed with 10 column volumes of binding buffer, and bound antibodies were eluted with elution buffer into microcentrifuge tubes containing 100µl 1M Tris pH 8.5 for neutralisation. Eluted fractions were analysed on a SDS-polyacrylamide gel for the presence of antibodies.

5.2.6.2 Covalent Binding of CYP6G1 for Antibody Purification

To purify the anti-CYP6G1 antibodies from other immunoglobulins, CYP6G1 was covalently bound to a column through which the antibody preparation was passed. Firstly, arginine was removed from purified, refolded CYP6G1 by concentration of the protein in a 10,000MWCO centrifuge filter concentrator (Millipore) and dilution in Ligand Coupling buffer (0.2M $NaHCO_3$, 0.5M NaCl, pH 8.3). A 1ml HiTrap NHS-activated HP column (GE Healthcare) was prepared following the manufacturers directions. Briefly, the column was injected with the CYP6G1 ligand solution and serially washed with NHS buffer A (0.5M ethanolamine, 0.5M NaCl, pH 8.3) and NHS buffer B (0.1M acetate, 0.5M NaCl, pH 4) to deactivate excess active groups.

Protein A column fractions containing antibodies were pooled and further diluted in 20ml PBS. The dilution was slowly (~1ml/min) passed through the CYP6G1-NHS column by syringe. Flow through was collected and the column was washed with 10ml PBS. 10µl samples were collected during the washing process to monitor the elution of serum proteins by SDS-PAGE. Bound antibodies were eluted with 8ml of 0.2M glycine pH 2.5, collected in 1ml fractions and immediately restored to pH 7 with 1M Tris pH 8.5. 10µl samples of the eluted fractions were analysed by SDS-PAGE, and fractions containing antibodies were pooled for use in Western blots of recombinant protein and *Cyp6g1*-expressing *D. melanogaster*. The antibodies were stored at 4°C.

5.2.6.3 Purification of CYP6G1 Antibodies using Antigen Bound to Nitrocellulose

In addition to the purification of antibodies using the target protein covalently bound to a column, an antibody adsorption method was used. A pellet of Rosetta cells expressing CYP6G1 was resuspended in PBS to an OD_{600} of 2.0 and from this 2x 1ml was pelleted. Each pellet was resuspended in 300µl of 1xSDS gel loading buffer, heated at 100°C and loaded onto an SDS gel prepared with a single well spanning the entire gel. Two gels were prepared, one for staining with Coomassie blue the other for blotting onto nitrocellulose for Western blotting.

The anti-CYP6G1 antibodies were purified following the affinity purification of monospecific antibodies using immobilised antigens method of Sambrook *et al.*, 1989. Briefly, the SDS-polyacrylamide gel was transferred onto a nitrocellulose filter (100v, 1h) and the filter was blocked with TPBS for 1h. 2ml of polyclonal antibody serum was added and kept at 4°C overnight. The antibody solution was discarded and the filter washed with 50ml 0.15M NaCl, then 50ml PBS. A scalpel was used to excise a strip of ~10mm from each side of the filter. The strips were incubated with gentle shaking in 50ml PBS with 40µl anti-rabbit antibody conjugated with alkaline phosphatase (Sigma) for 90min. The strips were stained with NBT-BCIP to visualise the bound antibodies (as described in Chapter 2). The strips were re-aligned with the major part of the filter and the position of the band corresponding to CYP6G1 marked and excised. The resulting anti-CYP6G1 antibodies were eluted with 250µl 0.2M glycine pH 2.5, 1mM EGTA. The eluate was collected into microcentrifuge tubes and

neutralised and stored as above, with 0.1 volume of 10x PBS and 0.02% sodium azide.

5.2.7 Western Blot using CYP6G1 Antibodies to Analyse Expression in *D. melanogaster*

The specificity of antibodies purified from the rabbit serum was analysed by Western blots, using crushed Hikone R and Canton S flies and recombinant CYP6G1. Ten male and ten female Hikone R flies, sacrificed by freezing, were crushed in 50µl PBS and diluted in 200µl 1x SDS gel loading buffer. After boiling at 100°C, 5min, the debris was pelleted by 10s microcentifugation at 16,000x g and 8µl was loaded onto a 10% SDS-polyacrylamide gel for Western blotting. Canton S flies were prepared in the same way. As a positive control, 1ml of an overnight-induced *Cyp6g1*-pET15b *E.coli* culture was pelleted and resuspended in 100µl 1x SDS gel loading buffer, from which 10µl was loaded onto the gel. Sperm samples from *D. melanogaster* (kindly provided by C. McCart) were also analysed using Western blots using the purified antibodies. The quantity of protein in each sample was measured using a EZQ Protein Quantitation Kit (Invitrogen) following manufacturers instructions. Western blots were performed as described in Chapter 2, using 20µl of antibodies purified with either of the techniques outlined, diluted in 50ml TPBS.

5.2.8 CYP6G1 Activity Assay

The activity of recombinant CYP6G1 was assessed using a methoxy-resorufin O-demethylation (MROD) fluorescence assay developed by Jenkins *et al.* (2006). CYP6G1 was used to supplement microsome preparations from Hikone-R and Canton-S flies. The reactions were prepared in a black flat-bottomed 96-well plate (Sterilin), containing 20µl of 25mM NADPH (Sigma), 30µl of protein preparation and 30µl fly microsomes, kindly provided by C. McCart. The reactions were made up to 190µl phosphate buffer (Na_2HPO_4 and NaH_2PO_4) at pH 7.8, and were started by the addition of 10µl 7-methoxy-3H-phenoxazin-3-one immediately prior to measuring optical activity with a Tecon Ultra fluorescence detector, set at 544nm excitation and 590nm emission. CYP6G1 was added as cleaved or uncleaved MBP fusion protein preparations, and as the lysate of cells expressing CYP6G1 with truncated or modified termini. Fluorescence was measured at 1-minute intervals for 1 hour.

5.2.9 Homology Modelling

A homology model was constructed for *D. melanogaster* CYP6G1. We were also interested in the interactions of CYP6G2, CYP6A2 and CYP12D1 with insecticides and models were created for these P450s using the same methods.

4.2.9.1 Selection of Reference Cytochrome P450 for Modelling

Sequence alignment of CYP6G1 was performed with the MUSCLE protein multiple sequence alignment software (www.drive5.com/muscle, Edgar, 2004). Sequences submitted to the alignment all came from the protein data bank (PDB, www.rcsb.org/pdb) and were chosen because structural information was available. The sequences were CYP2B4 (PDB reference number 1PO5), CYP2B4 with bound 4-(4-chlorophenyl)imidazole (1SUO), CYP2C5 (1DT6), CYP2C8 (1PQ2), CYP2C9 (1OG2), CYP3A4 (1TQN), CYP175A1 (1N97) and P450 BM-3 haem domain (1BU7).

5.2.9.2 Modelling

Homology models were made using the programme MOE (Chemical Computing Group, www.chemcomp.com). All of the sequences submitted to MUSCLE showed the greatest sequence identity with the human cytochrome P450 3A4, and the structure of CYP3A4 deposited by Eric Johnson's laboratory (Yano *et al.*, 2004) was used as the template for all of our models. CYP6G1 and CYP6G2 share 46.6% sequence identity and no additional model was made for CYP6G2.

Ten models were generated by MOE for each cytochrome P450, with the explicit inclusion of the haem coordinates available from the CYP3A4 structure, as used by Baudry *et al.* (2003). Each model was subjected to a coarse energy minimisation to remove possible van der Waals clashes between atoms. The best model for each cytochrome P450 was selected based on the MOE score output.

An oxohaem molecule was constructed with the MOE builder command and energy minimised with the CHARMm22 force field. The haem molecule was introduced into the homology models and a covalent bond was created between the haem iron and the sulphur of Cysteine 442 (CYP3A4 sequence). The haem-containing models were energy minimised as before, with iron geometry set to d2sp3 and oxygen geometry set to sp with charge set to zero. The programme "O" was used to locate errors in side-chain conformations and manually adjust uncommon covalent bond geometries in the models (Jones *et al.*, 1991). A final minimisation step was performed in MOE and the overall stereochemical quality of the models was assessed with the programme ProCheck (Laskowski *et al.*, 1993).

5.2.9.3 Substrate Docking and Cavity Calculations

Several compounds were docked into the cytochrome P450 homology models. The insecticides DDT, nitenpyram, diazinon and dicyclanil and the insect hormones ecdysone and juvenile hormone III were all created in the MOE builder command. These molecules were energy minimised using the MMFF94s force field and were roughly positioned into the active site cavities of the cytochrome P450 models using the MOE compute-simulations-docking procedure. For each ligand and enzyme model, one hundred possible conformations were generated, and in each case the model with the highest score in the MOE output was kept. The haem iron was set to a geometry of d2sp3 and the haem coordinates were fixed. The models were energy minimised using the MMFF94s force field, a dialectric of 1 and final energy gradient of 0.05 kcal/mol/Å. The difference between the potential energy of the entire minimised model and the sum of the individual potential energies for the protein and the ligand was used to calculate the interaction energy. The van der Waals and electrostatic components of the interaction energy were recorded for each cytochrome P450 and ligand. The cavity volume of the completed cytochrome P450 models was calculated with the VOIDOO package (Kleywegt & Jones, 1994). A probe of 1.4Å radius was used.

5.3 RESULTS

5.3.1 Cloning and Expression of CYP6G1-MBP

Cyp6g1 was cloned into parallel vectors to produce large amounts of recombinant protein for structural studies. Expression of both CYP6G1-MBP and -GST was detectable within 1h of induction, and the MBP construct consistently produced more protein. Increased yields came from overnight expression and substitution of LB with Terrific Broth.

Solubility of the fusion protein was estimated from SDS-PAGE of soluble and insoluble fractions isolated after cell lysis. Extensive trials were undertaken to improve the yield of soluble protein, and included the addition of the haem precursor δ-ALA, supplementation of cultures with 1μg/ml of antibiotics to induce cold-stress responses, 3% ethanol and a 1 min heat shock at 45°C to induce chaperone activity, and growth of cultures at a range of temperatures. Furthermore, chemical lysis using the BugBuster Protein Extraction kit (Novagen) and sonication were compared, with a range of lysis buffer pH and additive conditions. A maximum of approximately 50% of the CYP6G1-MBP fusion protein was detectable in the soluble phase after sonication, and came from EC100 cells grown in LB at 18°C. A 20mM Tris buffer adjusted to pH 7.0, with 0.5mM EDTA and 0.1mM DTT was used for sonication.

5.3.2 Purification and Cleavage of CYP6G1-MBP

CYP6G1-MBP was purified using amylose resin based affinity chromatography. The majority of soluble *E. coli* proteins released from cells by sonication did not bind to amylose in the purification column and were removed from the beads with two column volumes of a wash buffer. Introduction of 10mM maltose successfully eluted the bound proteins, which were revealed by SDS-PAGE to be approximately 110kDa. This corresponds to the 59.86kDa expected size of CYP6G1 plus the maltose binding protein and linker sequence. Fractions containing this protein were pooled for cleavage with the TEV recombinant protease. Cleavage was at maximum 50% complete, and separation of the cleaved CYP6G1 by passing through an amylose column yielded a very small quantity of protein that was difficult to detect by SDS-PAGE.

A batch purification method was also employed, allowing the protein to be traced at every stage. The soluble fraction derived from sonicated cells was added directly to amylose resin equilibrated in a no-salt lysis buffer at pH 7.0. After incubation on a rocker, the amylose beads were serially washed with TEV cleavage buffer and the amount of bound fusion protein was determined empirically by loading the beads directly onto an SDS-polyacrylamide gel. Approximately 80% or greater of the bound protein corresponded with the CYP6G1-MBP fusion protein, indicating a high level of purification was obtainable with this method (**Fig. 5.2 (A)**). The process was scaled up from a 100µl bed volume of amylose beads in a microcentrifuge tube by increasing the number of tubes or increasing the bed volume in a 50ml conical tube to accommodate the cell lysate from up to 3 litres of culture.

To cleave CYP6G1 from the amylose-bound MBP, TEV was added directly to the bead suspension and allowed to incubate on a rocker overnight. The sample was pelleted and the supernatant, containing cleaved CYP6G1, was removed by aspiration. The beads were washed with TEV buffer to remove residual protein, and the supernatant, bead wash and beads were analysed by SDS-PAGE. A further round of TEV protease addition was required, and the salt concentration of the TEV buffer was increased to reduce non-specific binding of proteins to the beads. Approximately 50% cleavage was achieved (**Fig. 5.2 (B)**). However, the CYP6G1 protein was not stable as a cleaved product and lost solubility without the MBP tag. We were unable to improve the cleavage efficiency, and the poor solubility of CYP6G1 meant that the large amounts required for structural studies could not be obtained.

A sample was prepared for dithionite-reduced and dithionite-reduced plus carbon monoxide difference spectroscopy, which confirms the presence of a haem group in the protein. A barely detectable hump was recorded at 450nm, which resulted from the very low level of CYP6G1 in the soluble phase, or from improperly folded protein that did not bind haem (data not shown).

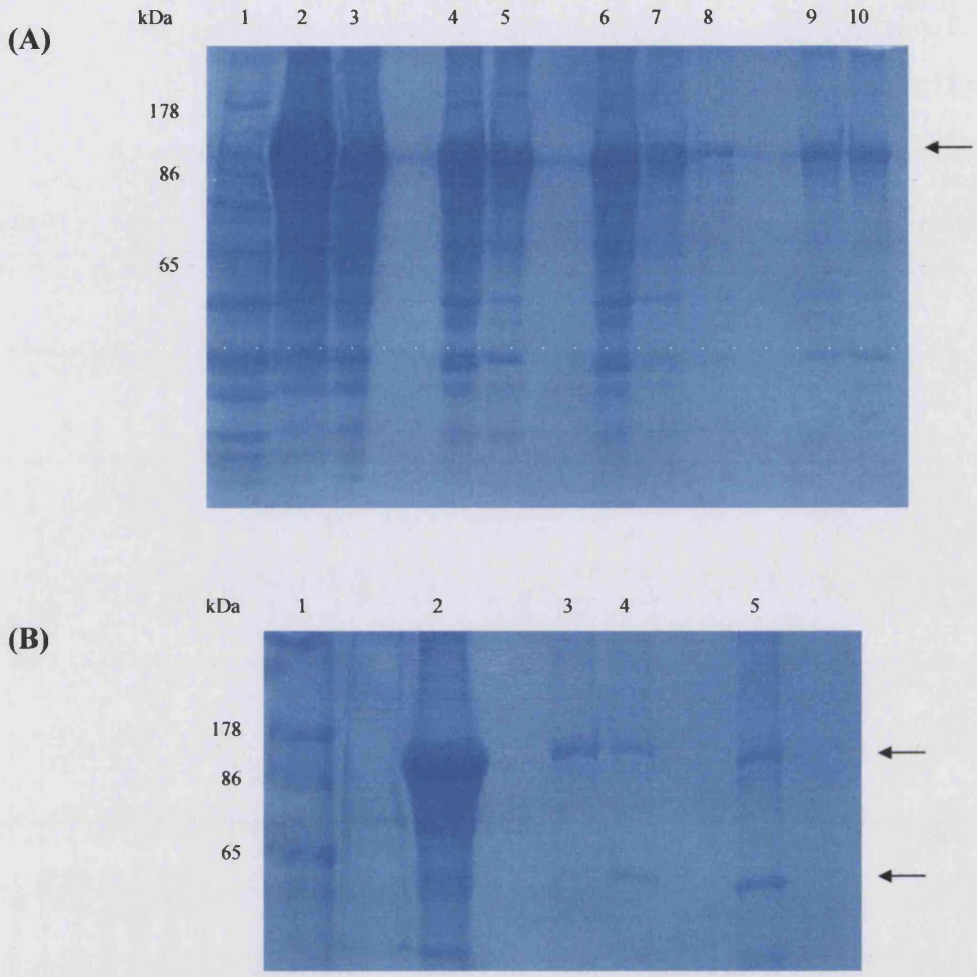


Fig. 5.2 CYP6G1 purification and fusion protein cleavage. **(A)** SDS-PAGE of CYP6G1 purified by batch method. Lane 1, uninduced culture; 2 and 3, culture induced for 24h; 4, soluble fraction; 5, insoluble fraction; 6, bead supernatant;, 7, wash 1 supernatant; 8, wash 2 supernatant; 9 and 10, amylose resin beads with bound CYP61-MBP (arrow). **(B)** Cleavage of CYP6G1, performed from batch purified fusion protein. The CYP6G1 was cleaved away from MBP still bound to the amylose resin beads. Lane 1, protein marker; 2, overexpression of CYP6G1-MBP after 24h; 3, resin with bound fusion protein (top arrow); 4, cleavage progress 1.5h after addition of TEV; 5, extent of cleavage after 19h incubation. Approximately 50% of the protein in the sample is cleaved, yielding CYP6G1 (bottom arrow).

5.3.3 Cloning of *Cyp6g1* into pET15b

To improve the solubility of recombinant CYP6G1, the gene was cloned into pET15b using a modification to the gene sequence. A putative transmembrane domain was identified at the N-terminal by analysis of the predicted amino acid structure using a ProtScale programme (ExPASy), and this region was deleted by designing primers to anneal downstream of the start codon. The first primer, designated Cyp6g1-T1 for 'truncated' N-terminus, was designed to remove the first 21 codons and introduce a new ATG start codon adjacent to the 22nd codon, producing a recombinant protein with increased hydrophilicity due to deletion of the transmembrane domain. A second primer, designated Cyp6g1-M1 for 'modified' N-terminus, was designed to replace the transmembrane domain with hydrophilic amino acids in the sequence MAKKTSSKG. This polypeptide chain has been as used previously to improve the solubility of the human cytochromes P450 2C8 and 2C9 (Schoch *et al.*, 2003, Williams *et al.*, 2003). A ProtScale plot of the predicted hydrophobicity shows an improvement to solubility with these modifications (**Fig. 5.3**). However, when expressed neither the truncated nor the modified proteins were soluble.

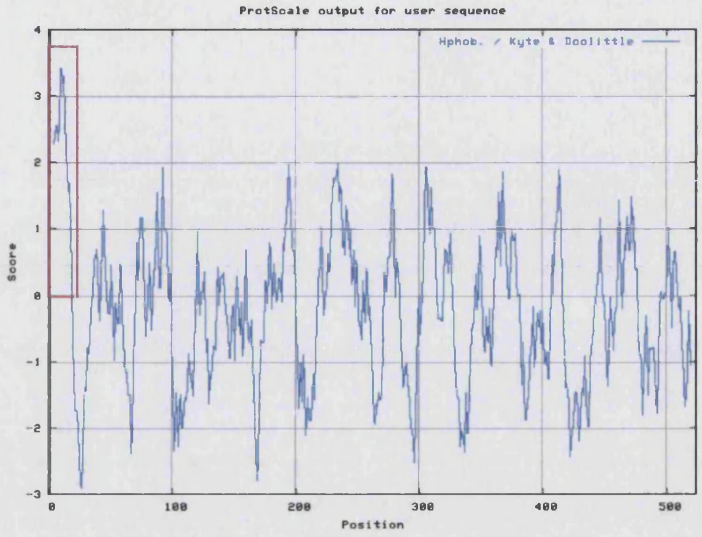
5.3.4 Use of Recombinant CYP6G1 to Purify Anti-CYP6G1 Antibodies

The poor solubility of CYP6G1 was prohibitive to structural studies, but the recombinant protein could be used for the small-scale purification of antibodies raised against a CYP6G1-MBP fusion protein.

5.3.4.1 Purification and Refolding of CYP6G1

Recombinant CYP6G1 was produced from the pET15b construct in DE3 Rosetta cells and released by sonication in a guanidine HCl buffer to denature the proteins. The pET15b vector in which the *Cyp6g1* gene was cloned introduced at the N-terminus a hexahistidine tag that allowed purification of the protein in a denatured state. The denatured lysate was pelleted and the soluble fraction was passed through a metal chelate membrane adsorber. CYP6G1 was retained by the membrane, which was thoroughly washed to remove unbound proteins. Solubilisation buffer containing 250mM imidazole was injected to elute the bound protein, but the presence of guanidine HCl in the eluate prevented analysis of the collected fractions by SDS-PAGE. All fractions collected after addition of imidazole were pooled and added dropwise to a stirred solution containing arginine to assist refolding. After overnight

(A)



(B)

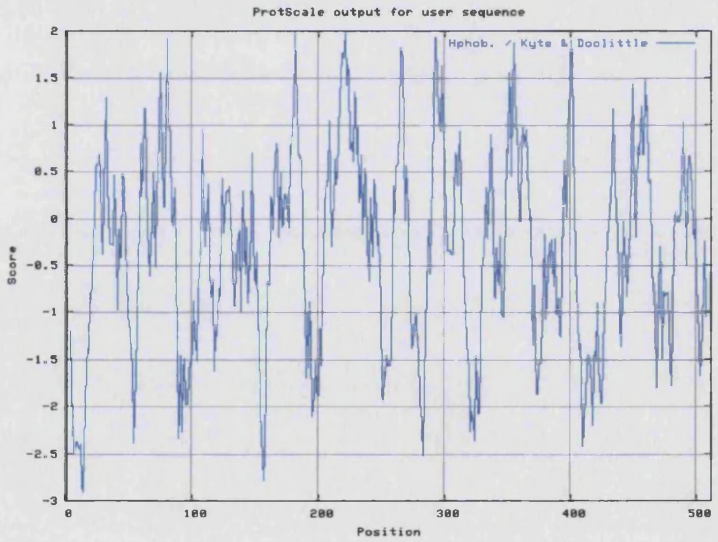


Fig. 5.3 Hydrophobicity profile of CYP6G1. Hydrophobicity plots of full-length (A) and N-terminally truncated (B) CYP6G1. The N-terminus of the full-length protein **(A)** has a region of ~20 highly hydrophobic residues and is predicted to form a transmembrane domain (red box), indicated by a large hydrophobic area on the hydrophobic side (positive values) of the midway line (0.0). Modification of the N-terminus through truncation of the transmembrane domain (not shown), or substitution with 8 hydrophilic residues **(B)** improves the solubility of the protein. The analysis was performed using ProtScale through ExPASy server, using Kyte and Doolittle values for amino acid hydrophobiocity and a window of 9 residues (Gasteiger *et al.*, 2005).

incubation the solution was concentrated and insoluble proteins were pelleted by centrifugation. SDS-polyacrylamide gel electrophoresis confirmed (a) that CYP6G1 was present in the metal chelate membrane elution, and (b) that the protein was soluble (**Fig. 5.4 (A)**). CYP6G1 was also of sufficient purity to prevent possible cross-reaction with the polyclonal antibody serum.

5.3.4.2 Purification of Anti-CYP6G1 Antibodies

Antibodies in the polyclonal rabbit serum were purified away from the serum proteins using a protein A column. Protein A derives from *Staphylococcus aureus* and has a high affinity for immunoglobulins from rabbits and other species (Hober *et al.*, 2007). Here, recombinant protein A covalently bound to sepharose was used as a ligand that captured the antibodies while the rabbit serum proteins were washed through. A low pH buffer was used to elute the bound antibodies, which were of sufficient concentration to be identified in elution fractions by SDS-PAGE (**Fig. 5.4 (B)**). The antibody elution was free of rabbit serum but contained immunoglobulins raised against CYP6G1 and MBP, as well as other native immunoglobulins. To reduce the level of cross-reactivity obtained in Western blots using these antibodies (not shown), a further purification step was used.

CYP6G1, purified with its hexahistidine tag under denaturing conditions and refolded in the presence of arginine, was loaded onto an NHS-activated column. The NHS column allows proteins to be covalently bound, so ligands that interact with the protein can be bound and eluted without releasing the coupled protein. Serum-free antibodies were passed though the column and antibodies not specific to CYP6G1 were washed through with PBS. Antibodies with an affinity for CYP6G1 bound under neutral conditions, but were eluted with a low pH elution buffer. Fractions containing these antibodies were pooled and used in Western blots to analyse the presence of CYP6G1 in Hikone-R and Canton-S *D. melanogaster*. The specificity of the antibodies was also examined, using *D. melanogaster* samples and recombinant protein, with particular comparison to the crude polyclonal antibody rabbit serum and antibodies purified using a second technique (described below).

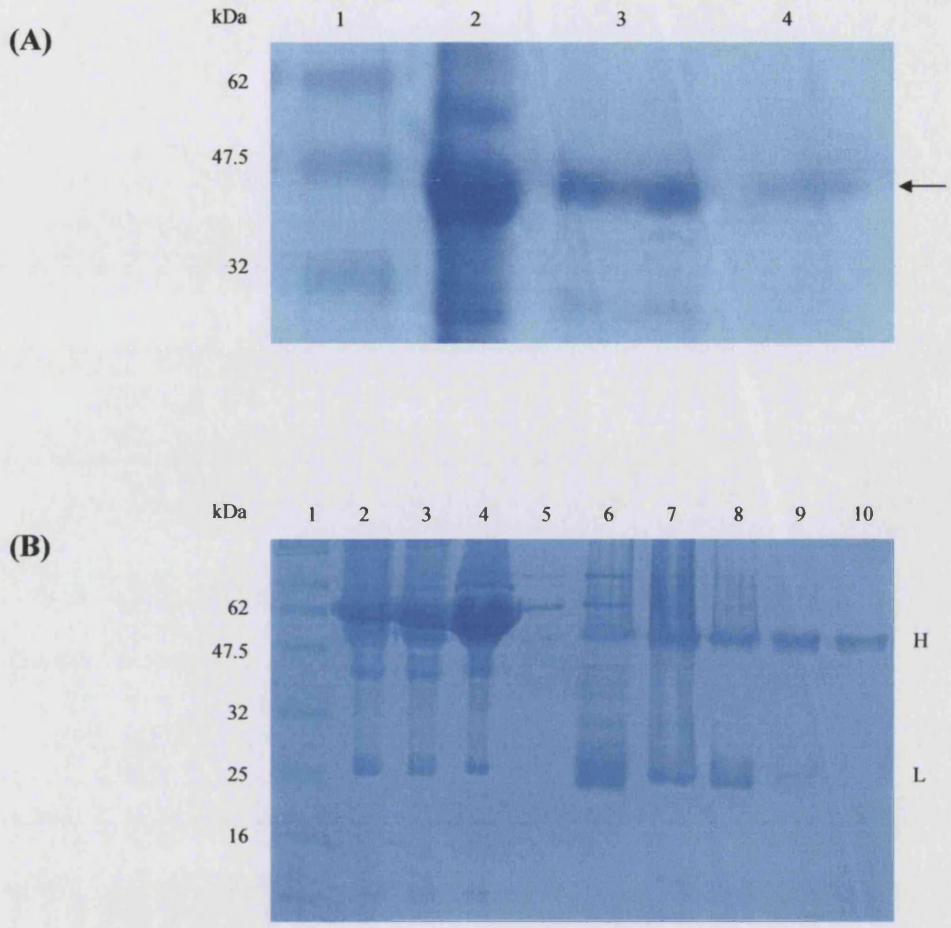


Fig. 5.4 Refolding CYP6G1 and purification of antibodies from serum. **(A)** *E. coli* expressing recombinant CYP6G1 were lysed and the proteins denatured with guanidine HCl. CYP6G1 was purified using a hexahistine tag at the N-terminus, and after elution from a metal chelate membrane was refolded by dropwise dilution into an arginine solution. The protein was concentrated and shown to be pure and soluble. Lane 1, protein marker; 2, whole cell extract; 3, soluble CYP6G1 (arrow) after refolding; 4, insoluble CYP6G1. **(B)** Polyclonal rabbit serum containing anti-CYP6G1 antibodies was diluted in 20mM sodium phosphate (pH 7.0) and loaded onto a Protein A column. After washing the column, bound antibodies were eluted with 100mM sodium citrate (pH 4.0). Eluted fractions were immediately brought to neutral pH to preserve their activity, and the anti-CYP6G1 IgGs were purified away from the other antibodies by passing through an NHS activated column to which refolded CYP6G1 was covalently bound. Lane 1, protein marker; 2, pre-column serum; 3, column flow through; 4, column wash, 5-10, eluted fractions. Note presence of serum-free heavy (H) and light (L) chains of the denatured antibodies.

A second method of purification separated anti-CYP6G1 antibodies from all other immunoglobulins using antigen bound to nitrocellulose. Recombinant CYP6G1 was heavily loaded on an SDS gel and transferred to a nitrocellulose filter. The filter was incubated with polyclonal antibody rabbit serum raised against CYP6G1-MBP. Narrow strips were removed from each end of the filter and bound immunoglobulins were visualised with an alkaline phosphatase anti-rabbit secondary antibody. A single large band was visible on these strips, as well as many small bands of cross-reacting immunoglobulins. The strips were re-aligned with the filter and the section corresponding to CYP6G1 was excised. Antibodies bound to this portion were eluted with a low pH buffer.

5.3.5 Expression of CYP6G1 in *D. melanogaster*

Western blots were performed as part of an analysis of CYP6G1 expression in *D. melanogaster*, and supported microscopic studies of sections of flies expressing GFP from the promoter of *Cyp6g1* (Chung *et al.*, 2007). Initially a crude sample of polyclonal antibodies from rabbit serum was used but produced high levels of cross-reactivity. The addition of 0.1% Tween in the antibody washes and 5% milk powder to block the membrane improved the blots but purification of the antibodies was necessary.

Two approaches were used and both techniques resulted in considerably reduced cross-reactivity. The amount of antibody yielded was low and experimentation with different dilutions was necessary to obtain blots with low cross-reactivity but a strong signal for the target protein. Recombinant protein loaded as a whole cell extract was used as a positive control, and verified the analysis of *D. melanogaster* samples. Comparisons between Hikone-R and Canton-S flies yielded the expected results, with more CYP6G1 identified in the resistant strain (**Fig. 5.5**). Antibodies purified from CYP6G1 covalently bound to a column, rather than immobilised on a nitrocellulose filter, were shown to produce the cleanest bands. Western blots were also performed on sperm extracted from Hikone-R and Canton-S. A signal was detected in both samples, but there was a difference between the level of CYP6G1 in these strains, indicating that male Hikone-R flies overexpress the protein in their sperm (**Fig 5.5**).

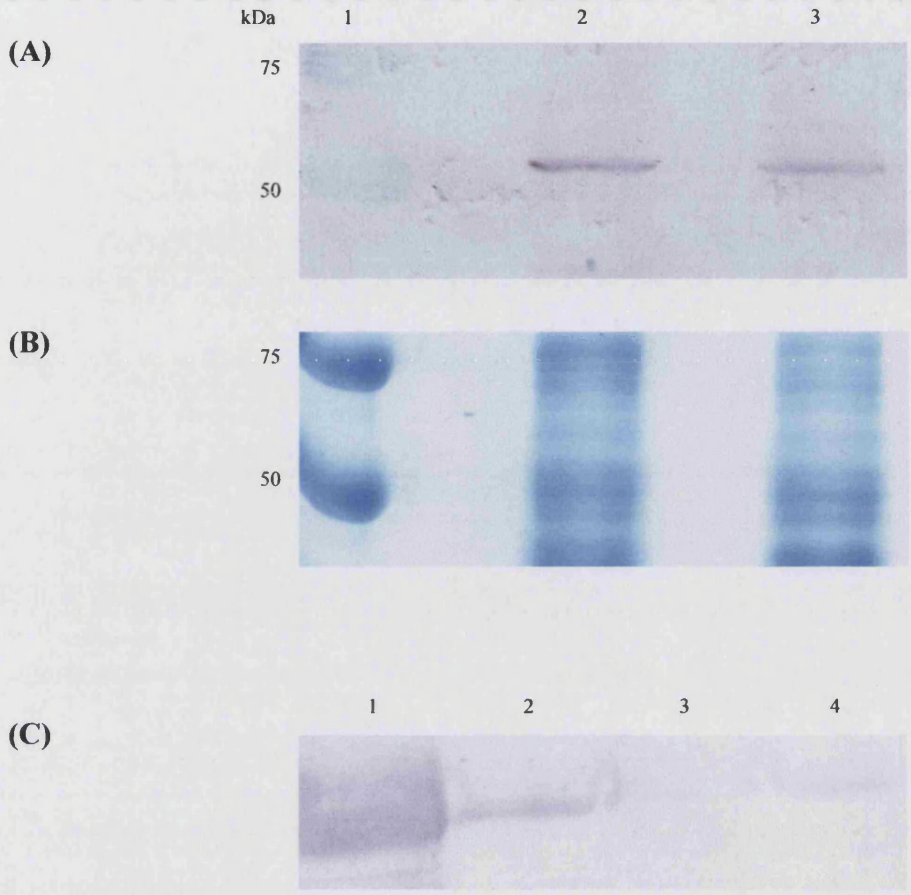


Fig. 5.5 Western blots showing CYP6G1 expression in *D. melanogaster*. **(A)** Western blot of CYP6G1 expression in total protein preparations of Hikone-R and Canton-S flies. Blot performed using purified anti-CYP6G1-MBP polyclonal antibodies shows greater expression of CYP6G1 the resistant flies, consistent with microarray data. Lane 1, protein marker; 2, Hikone-R total protein; 3, Canton-S total protein. **(B)** Coomassie stained SDS-PA gel of the same region, showing equal loading of fly samples. **(C)** Western blot showing hybridisation of the anti-CYP6G1 antibodies to recombinant CYP6G1 and to protein expressed in Hikone-R flies, Canton-S and Hikone-R sperm preparations. The blot shows that CYP6G1 is expressed in the sperm, and that higher levels are detected in the sperm of Hikone-R flies. Lane 1; CYP6G1 recombinant protein; 2, Hikone-R total protein; 3, Canton-S sperm; 4, Hikone-R sperm.

5.3.6 CYP6G1 Activity Assay

The O-demethylation of methoxy-resorufin ether by CYP6G1 was used to monitor activity of the recombinant protein. The product of the reaction, resorufin, is a fluorophore but the substrate is not. This reaction has previously been used to study the differential activity of microsomes prepared from Hikone-R and Canton-S flies, showing that the higher levels of CYP6G1 extracted from resistant flies correlates with greater production of the fluorophore (Jenkins *et al.*, 2006). Here, CYP6G1 expressed as a recombinant protein was unable to metabolise the O-demethylation of methoxy-resofurin in this experimental setup (data not shown), so the protein was added to fly microsome preparations to see if resorufin production could be increased. The addition of purified CYP6G1, with or without the MBP fusion, failed to affect the rate of MROD to the optically active resorufin product (**Fig. 5.6**). Furthermore, the lysate of cells containing CYP6G1 expressed from different constructs and with their N-termini truncated or modified also did not improve the rate of resorufin production. From these assays, it appears that the recombinant CYP6G1 is not biologically active.

5.3.7 Homology Modelling

Homology models were constructed for the *D. melanogaster* cytochromes P450 6G1, CYP12D1 and CYP6A2. An additional model for CYP6G2 was not created because it shares a high sequence identity with CYP6G1.

5.3.7.1 Sequence Alignment

The insect cytochrome P450 sequences were individually aligned against 8 reference sequences, chosen from the PDB because their structures have been solved (**Fig. 5.7**). In each case, the highest sequence identity came from the human cytochrome P450 3A4 (CYP6G1, 28.7%; CYP12D1, 23.7%; CYP6A2, 33.3%). The structure of CYP3A4 was therefore used as the template for all of our homology models. When the sequences of CYP6G1, CYP6G2, CYP12D1 and CYP6A2 were aligned with CYP3A4, several conserved regions were evident (**Fig. 5.8**). These correspond to motifs in helices K and I in the core of the protein, and a region of the haem-binding loop that contains the cysteine residue to which the haem group is covalently linked. The sequences Phe-X-X-Gly-X-Arg-X-Cys-X-Gly and Glu-X-X-Arg are absolutely conserved in cytochromes P450 (Werck-Reichhart & Feyereisen, 2000).

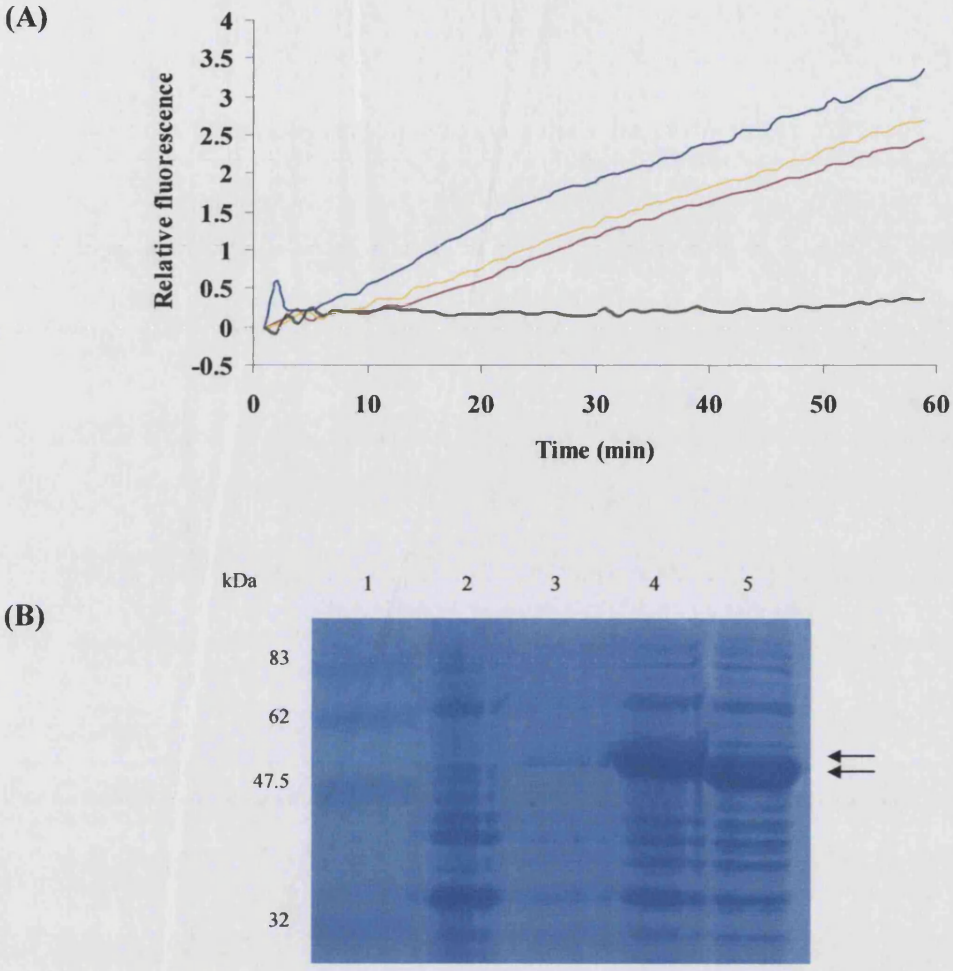


Fig. 5.6 CYP6G1 MROD assay. **(A)** The optical activity of resorufin, the product of methoxy-resorufin ether O-demethylation, was used to assess the biological activity of recombinant CYP6G1. Production of the fluorophore was measured over 1 hour, and shows no improvement to the reaction rate of CYP6G1 in Canton-S microsomes (blue) by the addition of purified CYP6G1. The yellow line shows fluorescence by a microsome preparation supplemented with uncleaved CYP6G1-MBP and red shows the addition of cleaved CYP6G1. The green line represents resorufin production by CYP6G1 with a transmembrane domain truncation, added to the reaction as unpurified cell lysate. Hikone-R microsomes produced considerably more resofurin in the same time period (not shown). **(B)** Expression of CYP6G1 with N-terminal modification (-M1, top arrow) and truncation (-T1, bottom arrow). Lane 1, protein marker; 2, uninduced CYP6G1-T1; 3, uninduced CYP6G1-M1; 4, induced CYP6G1-M1; 5, induced CYP6G1-T1.

(A)

Sequence	1	2	3	4	5	6	7	8	9
1: CYP6G1		16.6	17.6	17.8	19.0	20.2	27.6	13.8	17.8
2: 1PO5	17.1		74.8	38.1	38.0	35.6	18.4	14.5	13.0
3: 1SUO	18.3	74.8		47.7	48.8	43.8	18.8	14.5	13.4
4: 1DT6	17.8	36.8	46.0		65.9	65.1	16.9	16.9	14.3
5: 1PQ2	19.6	37.8	48.6	67.9		70.3	20.3	16.6	15.6
6: 1OG2	20.7	35.3	43.4	66.8	70.0		19.7	17.4	15.6
7: 1TQN	28.7	18.5	18.9	17.6	20.5	20.0		16.1	23.7
8: 1N97	11.8	12.0	14.0	14.5	13.8	14.5	13.2		16.3
9: 1BU7	18.0	12.7	13.1	14.5	15.3	15.4	23.1	19.2	

(B)

P450	Most favoured regions	Additional allowed regions	Generously allowed regions	Disallowed regions
CYP6G1	73.1%	24.1%	1.9%	0.9%
CYP12D1	74.0%	20.5%	2.6%	2.8%
CYP6A2	76.1%	20.4%	2.3%	1.2%

Fig. 5.7 P450 alignment scores and model Ramachandran plot statistics. **(A)** Alignment of CYP6G1 with P450 sequences selected from the Protein Data Bank shows greatest identity with TQN1, the sequence deposited for the human cytochrome P450 3A4. Alignments with the same sequences for *D. melanogaster* CYP12D1 and CYP6A2 also showed greatest identity with CYP3A4, with 23.7% and 33.3% identity respectively. All values shown are percentages. Alignment performed in MOE, and should be read using figures in the vertical column, e.g the alignment of CYP6G1 with CYP2B4 (1PO5) is 17.1%. **(B)** Ramachandran plot statistics for homology models of CYP6G1, CYP12D1 and CYP6A2. Greater sequence identity and energy minimisation resulted in fewer sterically disallowed conformations in the models for CYP6G1 and CYP6A2. However, for each model, >97% of residues were within generously allowed conformations.

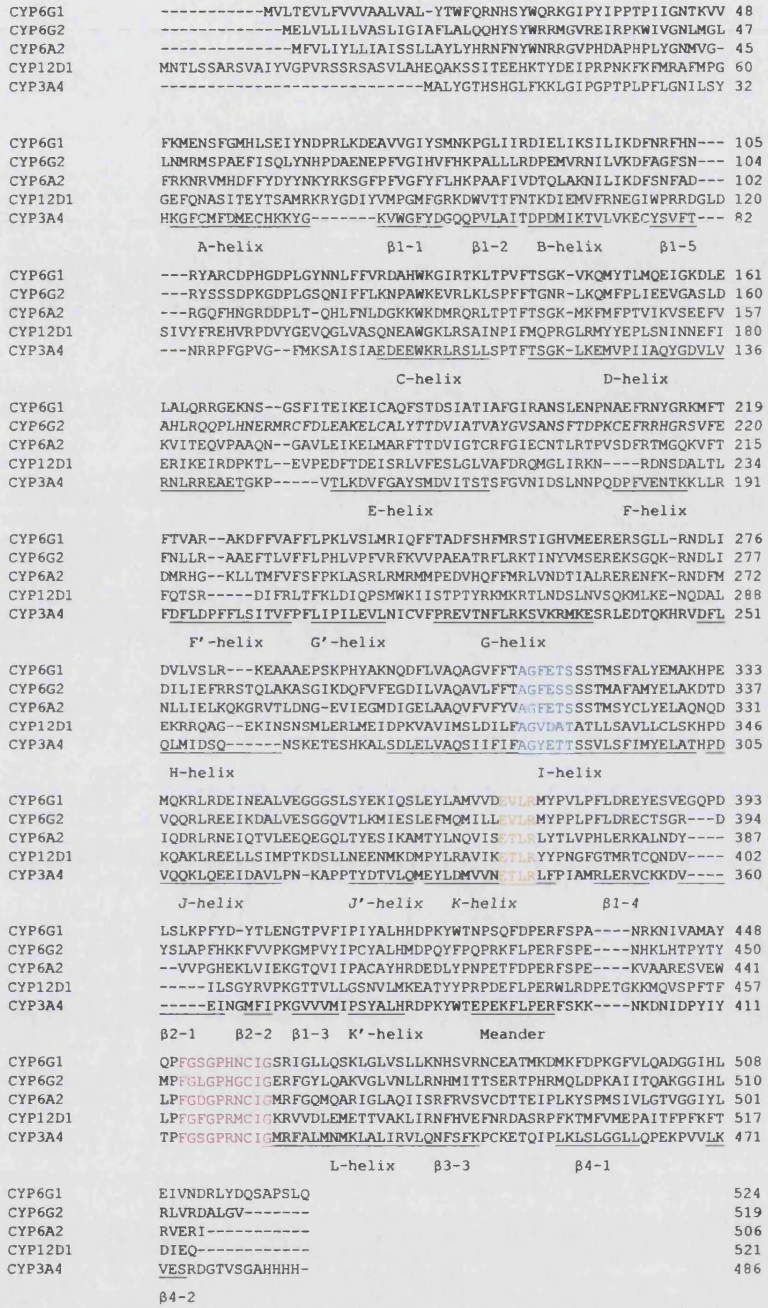


Fig. 5.8 Cytochrome P450 alignment. Alignment of CYP6G1, CYP6G2, CYP6A2, CYP12D1 and CYP2A4. Residues highly conserved among cytochromes P450 are highlighted. The Ala/Gly-Gly-X-Asp/Glu-Thr-Thr/Ser motif (blue) corresponds to the proton transfer groove on the distal side of the haem. The Glu-X-X-Arg motif (yellow) stabilises the core structure of the protein. Phe-X-X-Gly-X-Arg-X-Cys-X-Gly (red) is located on the proximal face of the haem just before the L helix.

5.3.7.2 Modelling

The homology models for CYP6G1, CYP12D1 and CYP6A2 were very similar in structure, but showed important differences in active site architecture. Generally, the models retained the overall protein fold of the template, CYP3A4, with a large domain comprised predominantly of α-helices, and a smaller, β-rich N-terminal domain (**Fig. 5.9**). The flexible F-G regions of CYP3A4 are poorly conserved among P450 sequences and structures, and consistently were poorly defined in our models. For instance, CYP6G1 has 6 additional residues in the F and F' helix region, resulting in loss of helix conformity and an extended loop region.

Residues that line the cavity of CYP6G1 come from elements found dispersively along the primary sequence. These include residues in helices F, G and I, as well as β-sheets 1-4 and 4-1. With the exception of the phenylalanine at position 133, all of the residues that appear to be available for interaction with substrates are located in substrate recognition sites (SRS) identified previously (Gotoh, 1992). Indeed, residues in all six recognised substrate recognition sites are in the enzyme cavity; SRS1 (Y107, P112, D111, F123), SRS2 (K216, F218), SRS3 (F245, F246), SRS4 (V308, T311, A312, E313, T316), SRS5 (V377, L378, P379) and SRS6 (V498, V499).

From our CYP12D1 model, it appears that SRS4, which resides in helix I, is heavily relied upon for substrate interactions, with residues E324, A325, G326, V327 and D328 all identified in the active site. SRS1 (R126, V141), SRS3 (K265) and SRS5 (N390, M395) are also implicated in substrate interactions. Consistently, residues predicted by sequence alignment to be in substrate recognition sites in CYP6A2 are found in the cavity of our model, including SRS1 (R103, N108, F119), SRS3 (M241), SRS4 (V309, A310, T314), SRS5 (V376, H378, L379, E380, R381) and SRS6 (V492).

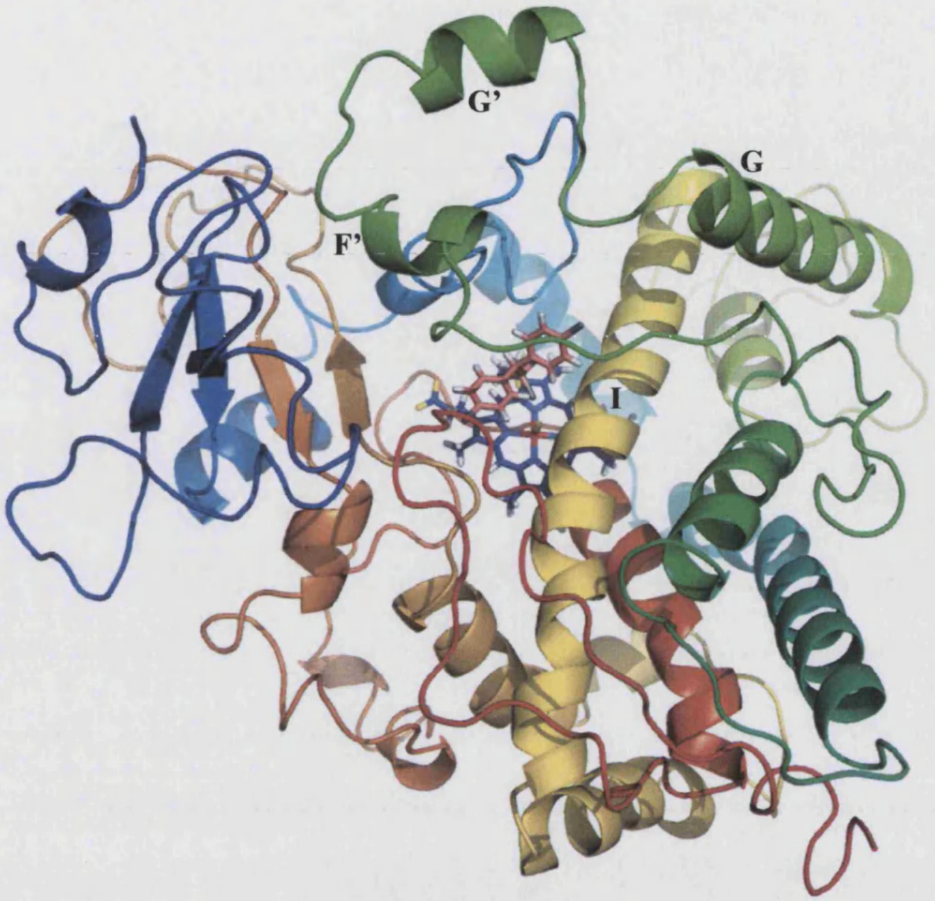


Fig. 5.9 Homology model of CYP6G1. Model created with MOE (Chemical Computing Group) using human CYP3A4 as a template. The protein shows the conserved overall fold of cytochromes P450, with a large helical domain and a smaller β-sheet domain. Colours shown as blue to red gradient from N-terminus to C-terminus. The large I helix spans the distal face of the haem group, and helices F' and G' are predicted to form a flexible 'lid' that closes over the active site cavity and allows substrate access. Haem and substrate (DDT) shown with atoms coloured red (iron), pale blue (nitrogen), light grey (hydrogen), dark grey (chlorine) and yellow (oxygen). Haem carbon atoms shown in blue and DDT carbon atoms shown in salmon for clarity. Image created in PyMol (DeLano Scientific).

5.3.7.3 Model Quality

The quality of the homology models was assessed with the programme ProCheck, which compares models with well-refined structures at the same resolution. Ramachandran plots are used to show the conformation values of phi and psi angles in the backbone of a polypeptide chain, and were prepared in ProCheck for each of the P450 models. For CYP6G1, CYP12D1 and CYP6A2, the Ramachandran plots showed that ~97% of all residues were within the generously allowed regions (**Fig. 5.7(B)**). In the case of CYP6G1 only 0.9% of residues were 'outside' of these regions. This corresponds to 4 residues (I41, L122, F250 and S291) that, because of their expected van der Waals radii, have sterically disallowed conformations when folded into the protein structure. The models for CYP12D1 and CYP6A2 had 2.8% and 1.2% of residues in disallowed conformations, respectively. These values are close to those of models published previously, such as the model for the insect cytochrome P450 6B1 (2% in generously allowed regions and 1% in disallowed regions (Baudry *et al.*, 2003)).

5.3.7.4 Substrate Docking

Substrate compounds were constructed in MOE builder and energy minimised, and then manually placed into the active site cavity of each P450 model. The substrates were docked using the MOE simulations docking procedure, and from the one hundred orientations explored, the substrate-protein model with the highest output score was selected for energy minimization. From the final model, van der Waals and electrostatic components of the interaction energy were calculated, as well as docking scores (**Table 5.1**).

Docking energies suggest that all four of the insecticides tested fitted into the cavity of CYP6G1. DDT, dicyclanil and nitenpyram had docking scores of -3.40, -2.63 and -3.43kcal/mol, respectively, and CYP6G1 has been shown previously to provide resistance to these compounds (Daborn *et al.*, 2007). Interaction energies suggest that, for CYP6G1, the most favourable interactions occur with nitenpyram. DDT was also show to form favourable interactions, with nine hydrophobic residues surrounding the insecticide. By contrast, our CYP6A2 model has just three hydrophobic residues surrounding the substrate, and the large size of the cavity means fewer contacts can be made (**Fig. 5.10**).

Cytochrome P450	Substrate	Relative resistance	Docking energy	Interaction energies		
				Total	van der Waals	Elect.
CYP6G1	DDT	4.06	-3.4001	-49.260	-29.343	-19.919
	Diazinon	-	-1.2100	-42.730	-25.621	-17.106
	Dicyclanil	2.23	-2.6300	-49.370	-10.443	-38.929
	Nitenpyram	1.96	-3.4293	-88.422	-9.313	-79.109
	JH III	n/a	-0.4	-39.9	-24.5	-25.3
	Ecdysone	n/a	26.3	-72.2	-19.3	-52.9
CYP12D1	DDT	2.36	-4.0736	-39.197	-24.621	-14.576
	Diazinon	-	-3.8265	-51.738	-15.282	-36.456
	Dicyclanil	1.48	-3.8563	-52.003	-9.474	-42.531
	Nitenpyram	-	-6.2013	-116.256	21.553	-137.809
CYP6A2	DDT	-	-7.0529	-35.181	-24.119	-11.062
	Diazinon	-	-3.8265	-49.502	-21.031	-28.471
	Dicyclanil	-	-3.2340	-29.121	-14.312	-13.889
	Nitenpyram	-	-5.0027	-37.846	-17.104	-20.742
CYP6G2	DDT	-	-3.4001	-49.260	-29.343	-19.919
	Diazinon	1.68	-1.2100	-42.730	-25.621	-17.106
	Dicyclanil	-	-2.6300	-49.370	-10.443	-38.929
	Nitenpyram	4.23	-3.4293	-88.422	-9.313	-79.109

Table 5.1 Relative resistance and docking energies. Biological data derive from individual cytochrome P450 genes overexpressed in *D. melanogaster* and are relative to the resistance of wild-type w^{1118} strain flies (from Daborn *et al.*, 2007). Docking energy, total interaction energy and its van der Waals and electrostatic energy components were calculated by MOE using the models we present (energies in kcal/mol). Note that CYP6G2 data are the same as for CYP6G1, as a separate model was not created. Boxes left blank (-) means relative resistance data were not determined.

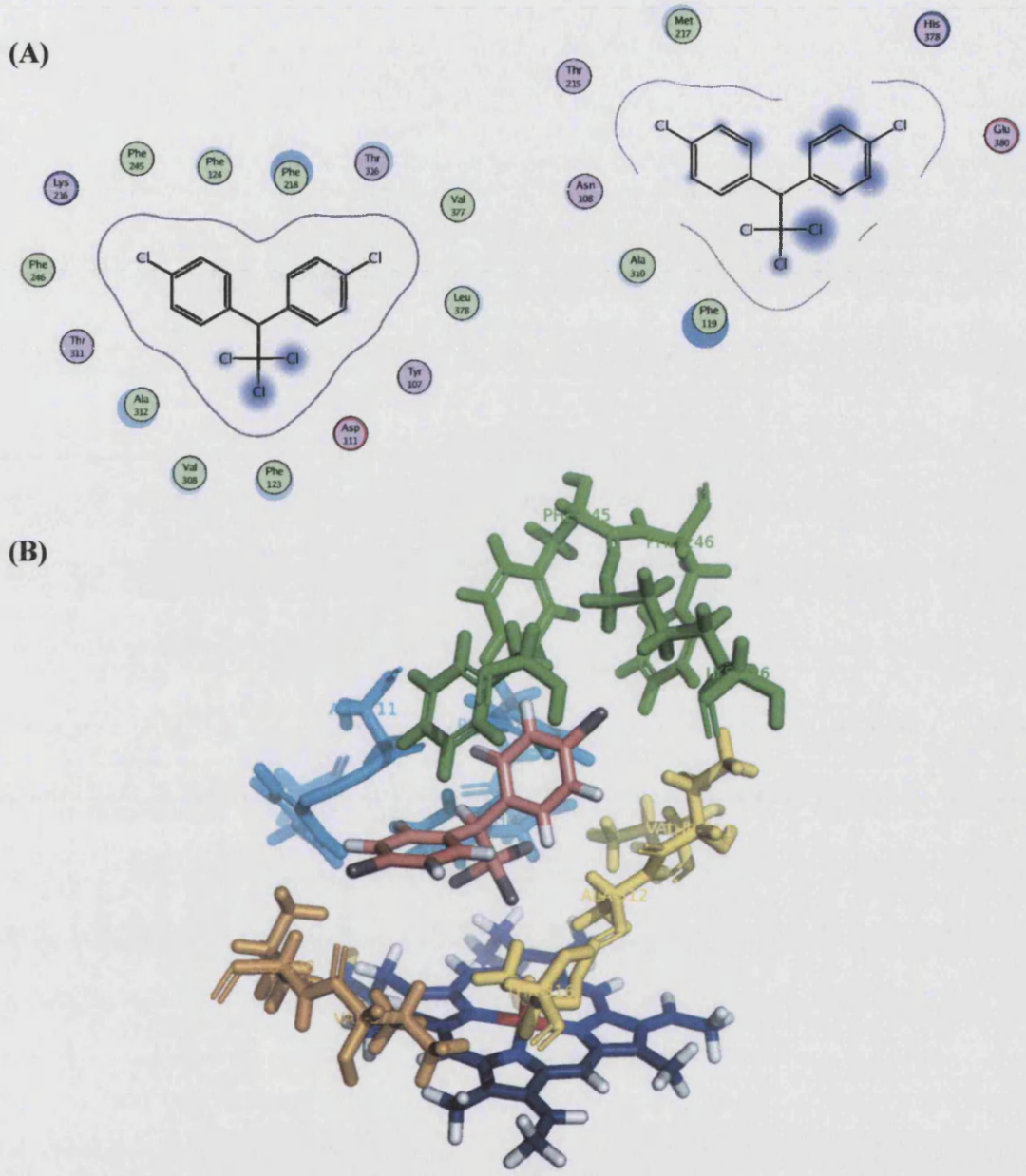


Fig. 5.10 Active site residues interacting with DDT. **(A)** CYP6G1 (left) residues are shown in close proximity to DDT in the closed active site of the protein. CYP6A2 (right) has a large cavity and fewer residues interact with the substrate, leaving large areas of solvent exposure. Residues in pink are polar and residues in green are hydrophobic. Red circle indicates a negative charge and blue circle indicates a positive charge. Blue shading around the residues and atoms of the substrate indicates solvent accessibility. Image created in MOE. **(B)** Rendered image of the same active site residues, showing their proximity to DDT. Colours and orientation the same as in **Fig. 5.9**. Image created in PyMol (DeLano Scientific).

Diazinon also fits into the cavity of CYP6G1, and positions well within the network of hydrophobic residues. The close proximity of tyrosine-107 and asparagine-111 to substrate carbon atoms, as well as phenylalanine-218 close to the substrate's only sulphur residue, likely explain the marginally less favourable interaction energy when compared to docked DDT. The docking energy for diazinon was also lower than for any of the other substrates tested, at -1.21kcal/mol. In the large cavity of CYP6A2, MOE identified just three residues in close proximity to diazinon that would make favourable contacts, but this also left a large proportion of the compound exposed to solvent (**Fig. 5.11**). When dicyclanil was docked into CYP6A2, potential interactions were observed between histidine-378 and arginine-103, but as with the other compounds tested, the substrate did not bind tightly and was largely solvent exposed.

Docking insecticidal compounds into the model of CYP12D1 resulted in docking scores of the order –3.8 to –6.2kcal/mol, reflecting the larger size of the active site cavity compared to CYP6G1. The most favourable interaction energy was obtained for nitenpyram. The organophosphorous compound diazinon and the insect growth regulator dicyclanil, although considerably different in shape and molecular weight, had almost identical docking energy and total interaction energy scores.

The insect hormones JH III and ecdysone were docked into our model of CYP6G1 to test the hypothesis that the enzyme may contribute to hormone titres in *D. melanogaster*. Ecdysone had a more favourable interaction energy, calculated from the van der Waals and electrostatic components, but the docking score was considerably higher than for any of the other substrates tested and the total complex energy was also unfavourable (484.1kcal/mol), strongly indicating that ecdysone does not fit in the active site cavity of the enzyme. The total complex energy for DDT and nitenpyram in CYP6G1 was –27.7 and 35.4kcal/mol, respectively, and JH III had a total complex energy of 23.2kcal/mol. Since biological data show that CYP6G1 can metabolise nitenpyram, these data suggest that the enzyme is also able to accommodate the insect hormone JH III with favourable docked energy.

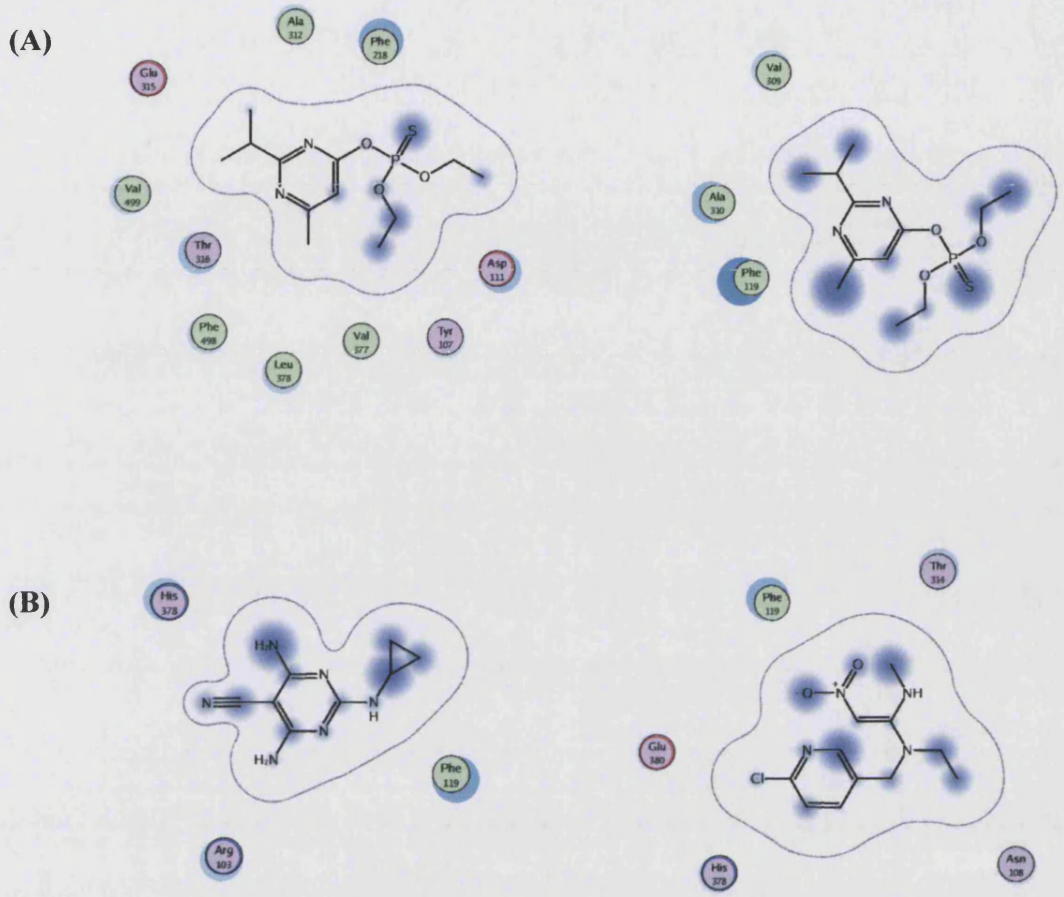


Fig. 5.11 Active site interactions with docked substrates. **(A)** Diazinon docked into CYP6G1 (left) and CYP6A2 (right). CYP6G1 has a smaller cavity than CYP6A2 and fits around the substrate well, but makes unfavourable interactions and had a poor interaction energy compared to the other insecticides tested. CYP6A2 makes very few interactions and is not associated with diazinon resistance. **(B)** Docking of CYP6A2 with dicyclanil (left) and nitenpyram (right). As with diazinon, CYP6A2 has a favourable docking energy because of the large cavity volume, but makes few interactions, particularly with dicyclanil, for which the least favourable interaction energy was measured out of all P450 and substrate combinations. Residues in pink are polar and residues in green are hydrophobic. Red circle indicates a negative charge and blue circle indicates a positive charge. Blue shading around the residues and atoms of the substrate indicates solvent accessibility. Images created in MOE (Chemical Computing Group) by Saskia Bakker.

5.3.7.5 Cavity Calculations

For each homology model the cavity volume was calculated using the programme VOIDOO. The cavity of CYP6G1 was calculated to be 72.63Å^3 and 'V' shaped, with the narrowest part of the cavity close to the haem group, and one branch of the 'V' extending towards the F-F' loop, a potential access channel. The cavity volume of CYP6A2 was much larger, and calculated to be 393.08Å^3. Similarly, the CYP12D1 model had a very large cavity. However, the cavity of CYP12D1 was characterised by VOIDOO as an invagination connected to the 'outside world', and calculation of its volume may have been distorted because its extent was not easily delineated. All of the volumes were considerably smaller than that of human CYP3A4, 1386Å^3, which was measured by Yano *et al.* (2004), using a 1.4Å probe in VOIDOO, as has been used here.

5.4 DISCUSSION

5.4.1 Expression and Purification of Recombinant CYP6G1

Expression of CYP6G1 as a fusion with the maltose binding protein yielded high levels of protein, but an initial analysis using an induction temperature of 37°C found that the majority of the protein was in the insoluble phase after cell lysis by sonication. A number of experimental trials were undertaken to increase solubility, both during expression and after harvesting the cells.

5.4.1.1 Adjustment to Growth Conditions

High levels of CYP6G1-MBP expression were obtained from Rosetta strain *E. coli* cells. The use of Terrific Broth also improved expression, but less protein was in the soluble phase in comparison to that produced from the EC100 strain of *E. coli* used for initial cloning. Similarly, the use of LB was marginally favoured for the expression of soluble protein. Low growth temperatures are frequently used to improve protein solubility by decreasing the rate of protein synthesis and altering the folding kinetics (Georgiou & Valax, 1996). Similarly, the lower expression rates in LB and EC100 are likely to have improved solubility. Consistently, we found a growth temperature of 18°C was optimal, and we used a low concentration of expression inducer throughout the trials.

In addition to a low growth temperature, Delcarte *et al.*, (2003) and others have allowed growth to high cell densities (OD_{600} >1.0) before induction and prolonged growth periods thereafter. The addition of δ-ALA as a haem precursor has been used widely in the production of recombinant cytochromes P450, ensuring the haem moiety will be present for activity and efficient folding of the protein. Although P450s are widespread in nature, *E. coli* does not contain any cytochrome P450 genes. However, haemoproteins are native to *E. coli* and the necessary machinery required for their synthesis is present (Kagawa *et al.*, 2003). *E. coli* has been the expression system of choice for the mammalian P450 structures that have been solved.

No chaperones were co-expressed with the *Cyp6g1*-vector constructs to assist in folding, but heat shocking to induce chaperone activity has been used elsewhere (Georgiou & Valax, 1996). Here, a trial culture was heated at 45°C immediately prior to induction, but was noted to have no observable effect. The addition of low levels of ethanol also induces the heat-shock response, and antibiotics that induce cold-stress responses have been used to enhance the heterologous expression of P450s (Kusano *et al.*, 1999), but they were shown to be ineffective for CYP6G1-MBP (not shown).

5.4.1.2 Adjustment to Post-Growth Conditions

Cells expressing CYP6G1-MBP were lysed in Tris at pH 7.4. This yielded marginally better results than lysis at pH 5.5 and 8.0, with fewer contaminating proteins in the soluble fractions. The dropwise addition of 0.1% Triton X or 0.01% SDS after centrifuging the sonication supernatant, in an attempt to remove the protein from any membranes, failed to improve solubility in any of the pH conditions. The salt concentration in the lysis buffer was also increased to 500mM as has been used elsewhere (Cosme *et al.*, 2003). Chemical lysis, as opposed to sonication, was tested but made no difference to the soluble CYP6G1 yield.

5.4.1.3 Purification and Cleavage of CYP6G1-MBP

Approximately 50% of the protein was in the soluble fraction after lysis, and the protein was purified with amylose resin. However, upon cleavage of the MBP tag with recombinant TEV protease, MBP could be observed by SDS-PAGE but very little CYP6G1 could be detected, indicating poor solubility. As a result, a sample prepared for UV difference spectroscopy failed give a spectrum.

After establishing the conditions at which CYP6G1-MBP could be eluted from amylose resin, and that the purity of the protein in these conditions was acceptable, a batch purification method was used which minimised loss of the protein. Cleavage in this system was also preferable, but neither process could easily be scaled up for large quantities, and the overall level of cleavage observed was still low, even after extended periods (overnight at 30°C). This possibly resulted from CYP6G1-MBP not being an efficient substrate for the protease, with conformations around the TEV site inhibiting protease access.

5.4.2 N-Terminal Modification of CYP6G1

For structural studies, and for the purification of CYP6G1 antibodies, large quantities of soluble protein were required. The CYP6G1 antibodies were raised against CYP6G1-MBP and provided in a polyclonal serum. To reduce background signal, we aimed to use a sample of CYP6G1 to purify the anti-CYP6G1 antibodies away from those raised against the MBP and other antibodies native to the serum. Since CYP6G1 cleaved away from its affinity tag was largely insoluble, a new strategy was taken; CYP6G1 was cloned into a pET15b vector, using modifications to the gene sequence used elsewhere to improve solubility.

Firstly, a construct was designed that removed a predicted N-terminal transmembrane domain. Such truncation has been used by Scott *et al.* (2001) and resulted in greater expression and solubility of mammalian CYP2B family proteins. A hydrophobicity plot of CYP6G1 suggests that the N-terminal domain appears to be the only membrane-spanning region. A similar conclusion has been made elsewhere for the mammalian P450 2C2 (Doray *et al.*, 2001). Deletion studies with this enzyme showed that the N-terminal region was the sole transmembrane domain and is both necessary and sufficient for membrane association. Furthermore, the presence of a polyhistidine tag at the N-terminus improved expression levels.

For a second construct, the hydrophilic sequence MAKKTSSKG was introduced in place of the 21-residue transmembrane region. This sequence has been successfully used by Schoch *et al.* (2004). Modification of the N-terminal region aims principally to improve the solubility of the protein, but also has the potential to improve expression because there is no longer the restriction imposed by the availability of internal membranes for sequestration of the recombinant protein (Scott *et al.*, 2001). A further modification has been used by Cosme and Johnson (2000) to create a soluble and monomeric protein when under high salt conditions. This involved substituting residues between helices F and G of CYP2C3 with corresponding residues from CYP2C5, which shows fewer tendencies to aggregate in reconstituted systems. Since it is unknown whether or not these residues are important to the activity of CYP6G1, such substitutions were not made here.

Expression trials with these constructs resulted in less soluble protein than was obtained from the CYP6G1-MBP fusion protein expression. The high solubility of MBP and the potential for the protein to display chaperone-like properties that promote folding is likely to have contributed to the solubility of CYP6G1 from our original construct (Kapust & Waugh, 1999). The insolubility of CYP6G1 produced from the pET15b vector construct was prohibitive to structural studies but did not impede the purification of anti-CYP6G1 antibodies. Because the protein did not need to be active, the potential loss of haem through refolding would not be a problem. CYP6G1 was solubilised in guanidine HCl and purified using the affinity of its N-terminal histidine residues for a nickel-based column. After elution, the protein was refolded in the presence of arginine to produce a low level of soluble protein that could be covalently bound to a column through which the anti-CYP6G1 polyclonal rabbit serum was passed.

The inability to produce large quantities of soluble protein remains a major barrier to the structural analysis of CYP6G1. As a consequence of eukaryotic cytochromes P450 being membranous and by nature highly hydrophobic, very few structures have been solved, and the majority of information about this large family of enzymes comes from the bacterial P450s. The few mammalian structures that have been solved provide suitable templates for homology modelling of other eukaryotic P450s.

5.4.3 CYP6G1 Activity Assay

The biological activity of the CYP6G1 produced here, from each of the constructs described, was monitored using a fluorescence assay. The O-demethylation of methoxy-resorufin ether can be catalysed by CYP6G1 in protein microsomes prepared from *D. melanogaster* (Jenkins *et al.*, 2006). The assay has been used to analyse differences in activity of CYP6G1 extracted from the midgut and Malpighian tubules of Hikone-R and Canton-S flies (McCart, 2006). Here, the assay failed to detect any resorufin production by purified recombinant CYP6G1, even in the presence of NADPH as an electron donor. Two other CYP6G1 preparations were tested, CYP6G1 with a truncated N-terminus and CYP6G1 with a modified N-terminus. Because these proteins are largely insoluble, they were not purified but added to the assay in cell lysate. Additionally, the assays were performed in the presence of fly

microsomes, in case the presence of an additional factor was necessary for optimal activity.

Low levels of CYP6G1 are expected in the microsomes of Canton-S flies, and were responsible for the production of the resorufin we observed. We found that the addition of recombinant CYP6G1 did not increase the rate of reaction, indicating that the protein was not biologically active. However, our results were not consistent, and with the individual supply of other cofactors such as NADPH P450 reductase it may be possible to detect some activity. Ensuring enough haem is available, through provision of δALA, is also necessary to optimise activity (Nishimoto *et al.*, 1993).

5.4.4 Homology Modelling

The structure of the cytochrome P450 6G1 from *D. melanogaster* was modelled to gain insights into its ability to efficiently metabolise DDT and to provide flies with cross-resistance to a number of novel insecticides. In addition, models were constructed for CYP6A2 and CYP12D1, and the structural differences between CYP6G1 and CYP6G2 have been considered. These cytochromes P450 were selected because they have been implicated previously in insecticide resistance (Amichot *et al.*, 2004, Brandt *et al.*, 2002) and recent transgenic fly experiments have provided comparable biological data for their metabolism of a range of synthetic compounds (Daborn *et al.*, 2007). Lastly, two insect hormones, juvenile hormone III and ecdysone were docked as a preliminary investigation into the ability of CYP6G1 to contribute to the hormone titres of *D. melanogaster*. All homology models presented here were based on the crystal structure of the human P450 3A4 (Yano *et al.*, 2004).

5.4.5 Crystal Structure of CYP3A4

The crystal structure of CYP3A4 was determined independently by Yano *et al.* (2004) at 2.05Å resolution, and by Williams *et al.*, (2004) at 2.8Å resolution. Like CYP6G1 in *Drosophila*, CYP3A4 is highly expressed and has a broad capacity to metabolise diverse substrates. It was chosen for structural analysis because it oxidises over 50% of the currently marketed pharmaceuticals and profoundly affects the efficacy of drug therapy. Furthermore, CYP3A4 is able to metabolise oestrogen, and an association has recently been made between oestrogen, the risk of breast cancer and the consumption of grapefruit, a known inhibitor of the enzyme (Monroe *et al.*, 2007).

CYP3A4 has the general tertiary fold identified in bacterial P450 monooxygenases, with a small N-terminal domain predominantly comprised of β-strands and a larger domain that is rich in α-helices and contains the haem and the active site (Williams *et al.*, 2004). The structure forms a large cavity that may be able to accommodate two or more substrates simultaneously and metabolise them without competitive inhibition (Scott & Halpert, 2005). This is possible because of the shape of the cavity, which has a relatively evenly distributed volume and is large close to the haem by comparison with the structure of CYP2C8 (Yano *et al.*, 2004).

The roof of the active site cavity is created by helices F and G. A lack of regular secondary structure in the residues that connect these helices to the short intervening helices F' and G' affords flexibility in this region. The F' helix forms an access channel with β-sheet 1, and helix G' forms a second channel that is separated from the first by a loop between helices B' and C (**Fig. 5.12**). It has been suggested that hydrophobic residues in these regions may interact with membranes, and that conformational changes necessary to invite large substrates may be mediated by a change in membrane properties or the binding of an electron transfer partner (Williams *et al.*, 2004).

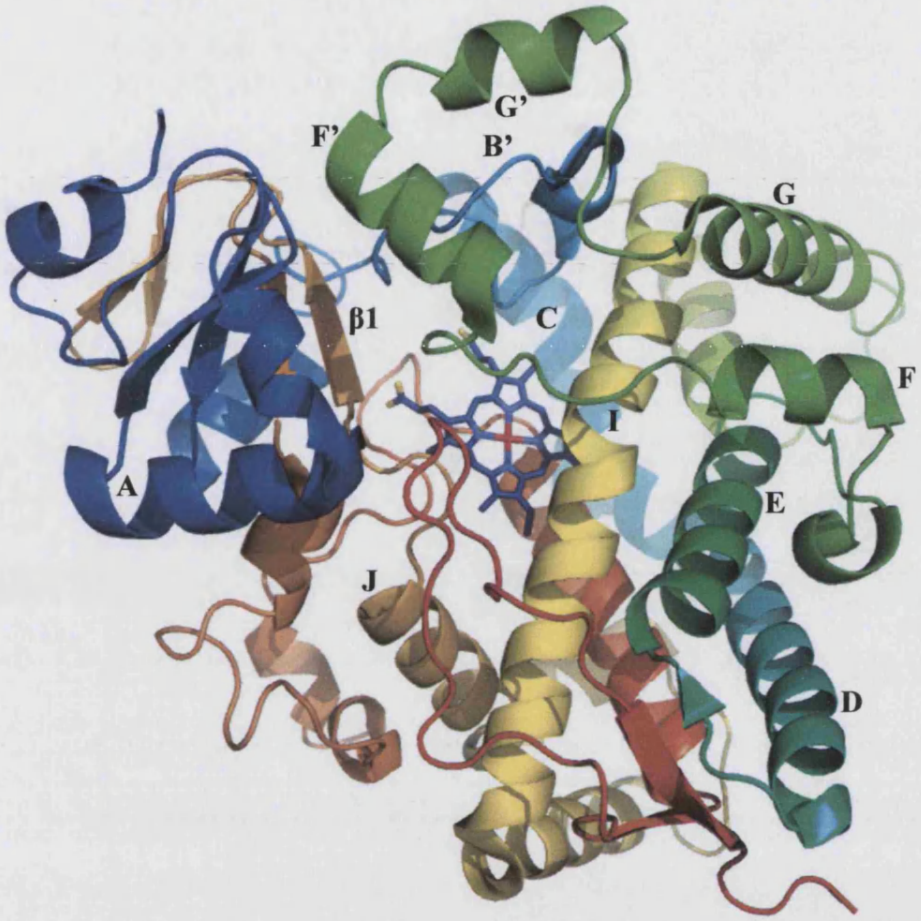


Fig. 5.12 Crystal structure of P450 3A4. Overall fold of CYP3A4 determined at 2.8Å resolution. The protein is coloured as blue to red gradient from N-terminus to C-terminus. Haem is shown with atoms coloured red (iron), pale blue (nitrogen), dark blue (carbon) light grey (hydrogen), dark grey (chlorine) and yellow (oxygen), viewed from the distal side and flanked by helix I. The flexible region F-G is visible as a lid that forms over the active site on the distal side of the haem. The loop between helix B' and helix C separates a channel formed between β-sheet 1 and the F' helix. Image constructed in PyMol (DeLano Scientific) using PDB 1TQN (Yano *et al.*, 2004).

5.4.6 Conserved Structure of Cytochromes P450

The structural fold of P450s is conserved despite greatly varying primary structures. Central to the role of these enzymes is the haem moiety, which is tethered to a universally conserved cysteine residue. The topology of the active site that surrounds the haem is diverse, and allows for the accommodation of substrates of significantly different sizes in different P450 members. Approximately 30% of residues in a P450 structure are generally well preserved within the superfamily, maintaining a deviation of less than 2Å (Mestres, 2005). These regions are mostly around the haem group and the proximal side of the protein, forming two clusters that include (i) sheets β1-2, β6-1, β1-3, helices B, K', K" and the cysteine pocket, and (ii) helices J and K and a four-helix bundle composed of parallel helices D, I and L, and the antiparallel helix E. Because of their conservation, these regions should produce accurate representations when homology models are created (Mestres, 2005). Other regions show greater structural flexibility, indicated by crystallographic *B*-factors, and have higher degrees of sequence and structural variability, including principally the surface helices B', F and G, and their adjacent structures helices C, C', H and the C-terminal part of helix I (Poulos, 2003, Mestres, 2005). These elements are distal to the haem, covering the active site, and are involved in substrate interactions.

Six sites in P450 structures appear to provide substrate recognition. These are dispersively located along the primary sequence, but lie along the access channel and cavity when the protein is folded, and predetermine the substrate specificity of each enzyme. The substrate recognition sites are in the B' helix, the F and G helix regions, helix I, and regions C-terminal to helices K and L (Gotoh, 1992). The position of these sites in flexible regions of the protein favours their role in substrate interaction. Furthermore, high sequence divergence in these regions is consistent with the ability of different members of the family to interact with different substrates. The active site of P450 enzymes is completely buried in the protein interior and is not connected to the outer molecular surface. Backbone movements and side-chain rotations contribute to large conformational changes that allow hydrophobic substrates to gain access to the haem site (Poulos, 2003).

5.4.7 Conformational Rearrangement and Access to the Active Site

Mammalian P450 2B4 forms an open structure with a cleft that divides the α-helical and β-rich domains, and extends from the outside surface of the protein to the haem site (Scott *et al.*, 2003). The cleft is approximately 15Å in width and is formed on one side by the F to G region and on the other side by the B'-C loop and helix C. The protein also forms a closed structure, apparently in response to entry and binding of a ligand in the active site (Scott *et al.*, 2004). The closed conformation is proposed to interact more favourably with the P450 redox partner, NADPH-cytochrome P450 reductase or cytochrome b_5, as a result of movements to helix C (Muralidhara *et al.*, 2006). In addition, closing the CYP2B4 structure creates more direct contacts between active site residues and the bound ligand that would not be possible if the protein remained open (Scott *et al.*, 2003).

The rearrangements within CYP2B4 that cause this closure are the largest that have been observed in a P450 (Scott *et al.*, 2004). The regions responsible correspond to regions that line the open cleft and, as described above, are the least conserved regions in P450 structures. Furthermore, the conformational changes appear to involve little secondary structural rearrangements, and the repositioning of tertiary elements that occurs on ligand binding does not disrupt the general P450 fold.

Being able to alternate between open and closed conformations allows diverse substrates to bind, but also allows stereo- and regiospecific metabolism once the substrate is docked at the active site (Scott *et al.*, 2004). Like CYP2B4, CYP2C5 forms a lid above the active site with helices F and G. This part of the protein is cantilevered over the top of helix I and rises and falls to allow substrate binding in the cavity (Wester *et al.*, 2003). The B' helix repositions depending on the substrate that binds, allowing for an induced fit conformational change that maximises van der Waals interactions with the substrate (Stout, 2004). Furthermore, the cavity of CYP2C5 contains a network of ordered water molecules that fill the space not occupied by the substrate. Hydrogen bonding between the water molecules and the substrate appears to be important for substrate specificity and is determined by the unique side chain composition of the cavity.

In P450cam, molecular dynamics simulations have been used to identify possible channels that allow access to and from the active site. Three different pathways were observed, the most likely of which is consistent with predictions for CYP2B4 and 2C5, and involves access through a crest-like channel that extends from helix I along the F-G loop and B' helix (Lüdemann *et al.*, 2000). Side-chain rotations and small backbone displacements of less than 2.4Å in these regions allow substrates to pass, while a second pathway requires greater flexibility and appears to be favoured by larger substrates. The simulations also show that a more hydrophilic channel, postulated from the crystal structure to be involved in substrate exit, is not used by ligands escaping from the active site.

5.4.8 Homology Model of CYP6G1

Our homology model of CYP6G1 conforms to the well-conserved structural fold of cytochromes P450. The cavity is buried inside the protein, but a potential substrate access channel exists above the pocket towards the F'-G' helix and loop region (**Fig. 5.13**). This is one branch in the 'V' shaped cavity, and is lined with hydrophobic residues. Given the 'breathing' motions identified in substrate-free and substrate-bound structures (Poulos, 2003), it is likely that movement of the F-G helix region occurs and allows the substrate to bind. This region also contains the substrate recognition sites delineated by Gotoh (1992). Our models for CYP12D1 and CYP6A2 were very similar to that of CYP6G1, except for the architecture and residue composition of the cavities. CYP6G1 has a less structured F' and G' region than CYP3A4, and in the case of CYP12D1 secondary structure is lost, resulting in an extended loop. Furthermore, the cavity of CYP12D1 reaches the outside molecular surface.

5.4.9 Homology Models and Insecticide Resistance

The docking energies of substrates in the active site cavities of the P450 models presented here can be compared with biological data. The ability of CYP6G1 to metabolise DDT is well established, but the ability of several other cytochromes P450 to metabolise this and other compounds has recently been assessed (Daborn *et al.*, 2007). For instance, compared with control flies, GAL4-UAS transgenic flies overexpressing *Cyp6g1* show a relative resistance of 4.06 to DDT, 1.96 to nitenpyram and 2.23 to dicyclanil. Overexpression of *Cyp12d1* results in a relative resistance of

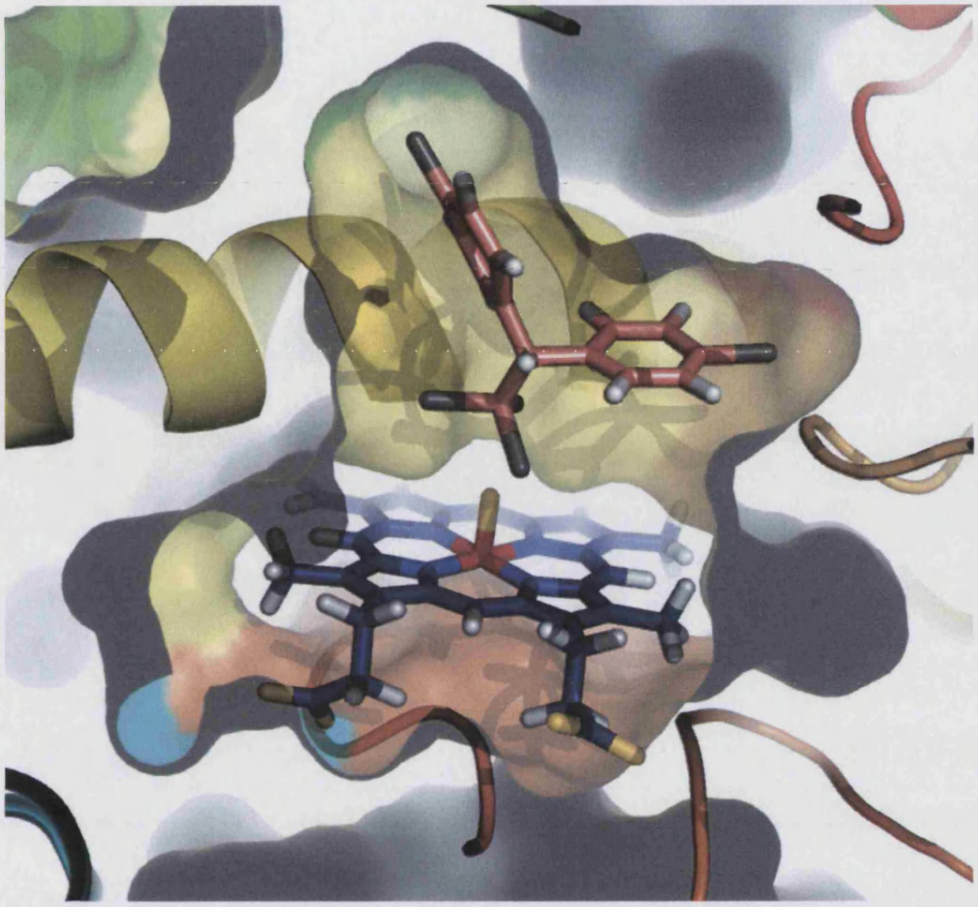


Fig. 5.13 CYP6G1 active site. DDT is shown to fit tightly inside the 'V' shaped cavity of CYP6G1. The substrate is in close proximity to the I helix and interacts with residues Thr 316, Ala 312 and Val 308. The cavity is buried inside the protein, but a likely access channel exists where the cavity extends at the top of the image, towards the F' and G' helices. The closest chlorine atom to the haem oxygen is 3.12Å away, calculated with MOE (Chemical Computing Group). Image created in PyMol (DeLano Scientific).

2.36 to DDT and 1.48 to dicyclanil. Unfortunately, the relative resistance ratios of some enzyme and substrate combinations were not determined, preventing direct comparisons between, for example, the metabolism of nitenpyram by CYP6G1 and CYP12D1.

5.4.9.1 Insecticide Resistance Conferred by CYP6G1

Of the eight cytochrome P450 genes tested by Daborn *et al.* (2007) the highest level of resistance to DDT was recorded by overexpression of *Cyp6g1*. Our model shows that the insecticide fits well into the cavity of the enzyme, which is replete with phenylalanine and other hydrophobic residues. DDT is highly hydrophobic, and has five chlorine atoms that are largely surrounded by the hydrophobic side chains. Accordingly, for DDT both components of the interaction energy are more favourable for CYP6G1 than with the other P450s modelled.

CYP6G1 is not restricted to the metabolism of DDT and has been implicated in conferring cross-resistance to a number of insecticides (Daborn *et al.*, 2002). Consistently, we found that nitenpyram and dicyclanil have favourable docking scores and interaction energies. The least favourable values were observed for diazinon, an insecticide apparently not metabolised by CYP6G1 when tested by Daborn *et al.* (2007), although resistance to diazinon has previously been associated with the enzyme (Pyke *et al.*, 2004). Diazinon is large and polar, and forms some unfavourable interactions in the active site; e.g., several polar residues are in close proximity to hydrophobic regions of the substrate.

5.4.9.2 Insecticide Resistance Conferred by CYP6A2

Daborn *et al.* (2007) noted that no increased survival was detected on any tested insecticide by overexpression of *Cyp6a2*. It can be assumed that CYP6A2 does not metabolise DDT, nitenpyram, diazinon or dicyclanil, or that these compounds are metabolised to toxic products. Our model of CYP6A2 has a very large cavity. This would accommodate even the largest of the substrates, as indicated by the low docking scores, but does not allow a tight fit and much of the substrate remains solvent-exposed. Furthermore, large cavities allow the possibility of non-productive as well as productive orientations to be occupied.

Amichot *et al.*, (2004) found that resistance to DDT conferred by CYP6A2 was associated with three point mutations; R335S, L336V and V476L. The gene sequence used in the present study and by Daborn *et al.*, (2007) derives from a *y; cn bw sp* strain of flies, and the CYP6A2wt protein from these flies shows a low capacity to metabolise DDT when expressed heterologously. However, heterologous expression of the mutant protein (CYP6A2vSVL) that provides resistance in the DDT^R strain shows a high level of DDT metabolism (Amichot *et al.*, 2004). Interestingly, the three point mutations associated with resistance are not in the active site of CYP6A2 but in helix J and at the limit of the β3-3 sheet. These residues are located at distance to those believed to be involved in substrate binding, and the mutations do not appear to interfere with substrate affinity.

It is not clear how these subtle changes to the protein have enabled it to increase its metabolism of DDT. The arginine at position 335 appears to form a salt bridge with aspartic acid at position 331. Substitution to serine will prevent the formation of this contact and affect the local stability. Consistently, Amichot *et al.* (2004) noted that the mutant form is less stable than the wild-type protein. Because the ratio of the different metabolites of DDT differs between CYP6A2wt and CYP6A2vSVL, the mutations may allow a difference in substrate orientation within the active site (Amichot *et al.*, 2004).

Improvements to catalytic activity by mutations outside of the active site have been reported previously. The cytochrome P450 BM-3 has been converted by laboratory evolution from a fatty acid monooxygenase into an efficient catalyst of the hydroxylation of alkanes to alcohol (Glieder *et al.*, 2002). Only one of the eleven mutations identified in the most efficient enzyme produced, BM3 139-9, was in the active site and in close contact with the substrate. Five of the mutations were in helices F and G, and the loop that connects them. This is the region that undergoes the largest conformational change and acts as a lid that closes over the substrate access channel upon substrate binding.

Mutation in helix J has also been described for the improvement of P450 catalytic activity. In the mammalian cytochrome P450 2B1, introduction of the S334P mutation enhanced the catalytic efficiency of all of the substrates tested (Kumar *et al.*, 2005). The highest rate of activity was achieved in a triple mutant; the second mutation was in helix F in a region important in substrate recognition, and the third was in helix E. The authors suggest that the changes in helices E and J may exert long-range effects, with the possibility that the amino acids involved are key residues in the network of interactions that extend throughout the protein.

5.4.9.3 Insecticide Resistance Conferred by CYP12D1

The gene *Cyp12d1* is overexpressed in the DDT resistant strain of *D. melanogaster* *Rst(2)DDT*^{Wisconsin} (Brandt *et al.*, 2002). *Cyp6g1* is also overexpressed in these flies, and both are likely to contribute to DDT resistance. Consistently, Daborn *et al.* (2007) found that *Cyp12d1* overexpression lines were resistant to DDT, and these flies also showed a relative resistance to dicyclanil. The large cavity of CYP12D1 has yielded favourable docking scores for all of the insecticides tested here. For instance, the energy value obtained for docking DDT was –4.07kcal/mol for CYP12D1 and –3.40kcal/mol for CYP6G1, indicating more freedom of movement for the substrate in the former. However, the interaction energy was more favourable for CYP6G1, which has fewer hydrophilic residues in the active site and is able to make appropriate interactions with the highly hydrophobic compound. The most notable interaction energy value came from docking nitenpyram, with an apparently highly favourable energy of –116.26kcal/mol. CYP12D1 has not been implicated in resistance to nitenpyram, and this incongruity may result from inaccuracies in our model; 2.8% of residues were in disallowed conformations and the van der Waals component of the interaction energy was high. This suggests that there are more steric repulsions in the interaction with nitenpyram than with the other substrates. The inaccuracy of the CYP12D1 model is also evident in the docking of diazinon and dicyclanil. These substrates fit with almost identical docking scores and interaction energies, but biological data suggest only dicyclanil can be metabolised to provide resistance (Daborn *et al.*, 2007).

5.4.9.4 Insecticide Resistance Conferred by CYP6G2

No homology model was made for CYP6G2, as a consequence of the high amino acid sequence identity it shares with CYP6G1, the main focus of this study. However, transgenic flies overexpressing *Cyp6g2* were found to have a different metabolic profile to those overexpressing *Cyp6g1* (Daborn *et al.*, 2007). CYP6G1 did not confer resistance to diazinon, while CYP6G2 showed a relative resistance to this insecticide of 1.68 over control flies. Similarly, a higher level of resistance to nitenpyram was recorded by GAL4-UAS *Cyp6g2* flies in comparison to GAL4-UAS *Cyp6g1* flies (relative resistance of 4.23 and 1.96, respectively). As analysis of CYP6A2 has shown (above), mutations need not be in the active site to exert an affect on the metabolic activity of a cytochrome P450. However, our model predicts that there are some residue differences between the active sites of these two proteins, and these residues are the best candidates to explain the different activities towards nitenpyram and diazinon. Relative resistance values for DDT and dicyclanil are not available for *Cyp6g2*-overexpressing lines.

Presuming the structures of CYP6G1 and CYP6G2 are very similar, and that the spatial orientation of elements that interact with the substrate are conserved, eight residues that line the inside of the cavity differ between the two enzymes. Five of these involve substitutions between small hydrophobic residues (F245V, F246V, V308L, F498I and V499I (numbering according to CYP6G1 sequence)) and may induce little change except increase the net size of the cavity. The substitution T316S is also expected to induce little change in the active site. The most significant substitution is likely to be that at residue 216; in CYP6G1 this is a lysine, and in CYP6G2 this is a serine, a substitution that results in the loss of local positive charge. Lysine-216 is in the F-F' loop of CYP6G1. When DDT occupies the active site, this residue is in close proximity to a chlorine atom and is positioned directly above the substrate, marking the extent of the internal cavity. Lastly, valine-377 is a proline in CYP6G2. Our homology model shows that this residue is between helix K and β-strand 1-4, and the proline is therefore not located at a site where it can interrupt an α-helix. However, like the other residues discussed, V377 resides in a substrate recognition site, and therefore has the potential to alter the metabolic profile of the enzyme, so reducing its interaction with DDT, improving its ability to metabolise

nitenpyram, and introducing a capacity to metabolise the organophosphorous compound diazinon.

Cyp6g2 had not been implicated in insecticide resistance but was selected by Daborn *et al.* (2007) due to its location on the same gene cluster in *D. melanogaster* as *Cyp6g1*, and because the two genes are close paralogs. Resistance of a wild-caught strain of *D. melanogaster*, referred to as *Innisfail*, to the insecticide diazinon was shown to be necessarily multi-factorial (Pyke *et al.*, 2004). One factor mapped to a locus on chromosome III and the other to a region between the loci *en* and *sca* of chromosome II. Given the close proximity of *Cyp6g1* and *Cyp6g2* in this region, and the observed resistance to diazinon reported by Daborn *et al.* (2007), the possibility should not be ignored that the resistance displayed by *Innisfail* flies was mediated by *Cyp6g2*. The identification of an *Accord* insertion upstream of *Cyp6g1*, however, makes CYP6G1-mediated resistance the most parsimonious explanation (Pyke *et al.*, 2004).

5.4.10 Expression of *Cyp6g1* in *D. melanogaster*

The ability of CYP6G1 to provide a high level of insecticide resistance in *D. melanogaster* is promoted by its tissue-specific expression. It has recently been shown that the *Accord* element inserted 291bp upstream of the *Cyp6g1* start site in Hikone-R drives overexpression in proportion to basal expression levels in different tissues (Chung *et al.*, 2007). This results in high levels of CYP6G1 in the gastric cecum, midgut, Malpighian tubules and fat body, all tissues of the insect associated with detoxification activities. Indeed, the Malpighian tubules appear to be the limiting tissue in control of DDT-induced mortality, and expression of CYP6G1 in the tubules provides the baseline protection against DDT in adult flies (Yang *et al.*, 2007).

Cyp6g1 overexpression is also controlled temporally, with low levels detected in the embryos and high levels in larvae. Since imidacloprid acts as a larvicide, and Hikone-R *Drosophila* are not susceptible to this compound, larval expression of *Cyp6g1* is not unexpected. Although Hikone-R larvae produce more *Cyp6g1* transcript than do Canton-S, it is interesting to note that both resistant and susceptible larvae produce high levels, suggesting a possible role of the gene in this stage of development (McCart, 2006). In adult flies, resistant genotypes are associated with up to 100-fold

greater levels of *Cyp6g1* transcript. Flies carrying the resistant allele are more fecund than homozygous susceptible flies, and show accelerated developmental rates at larval and pupal stages (McCart *et al.*, 2005). Furthermore, egg, larval and pupal insects with the resistant genotype are more viable than their susceptible counterparts and this occurrence is more pronounced when the resistance gene comes from the female parent, suggesting a possible maternal effect. This hypothesis is supported by an observation of higher *Cyp6g1* transcript levels in the embryos of these flies.

The female sperm storage organ (spermatheca) of transgenic flies expressing GFP from the *Cyp6g1* promoter region using the GAL4 UAS system, has been shown to fluoresce (McCart, 2006). Crosses using the strains 'paired' and 'son of Tudor' showed that this fluorescence was associated not with expression in the spermatheca induced by the presence of Hikone-R sperm or accessory proteins, but was associated with the sperm itself. In the current study, we used purified anti-CYP6G1 antibodies to confirm that the signal observed was not auto-fluorescence and show (a) that CYP6G1 is present in sperm, and (b) that the sperm of Hikone-R flies contains more CYP6G1 than does that of Canton-S sperm.

5.4.11 Role of CYP6G1 in Reproduction

Recombinant CYP6G1 was covalently bound to a purification column, through which diluted anti-CYP6G1 polyclonal serum was passed. Eluted antibodies were used in a Western blot prepared with the sperm of Hikone-R and Canton-S flies. As expected, CYP6G1 was detectable in each sperm sample, but the Hikone-R sperm gave a stronger signal. Given that Canton-S sperm also expresses CYP6G1, and the relatively low level of protein that can be produced in comparison to tissues of the alimentary canal and excretory system, it is unlikely that CYP6G1 in the sperm has a role in insecticide resistance. Lastly, it should be noted that the GAL4 GFP system also revealed low levels of CYP6G1 expression in the female reproductive tissues, further indicating a role of the protein in reproduction (**Fig. 5.14**).

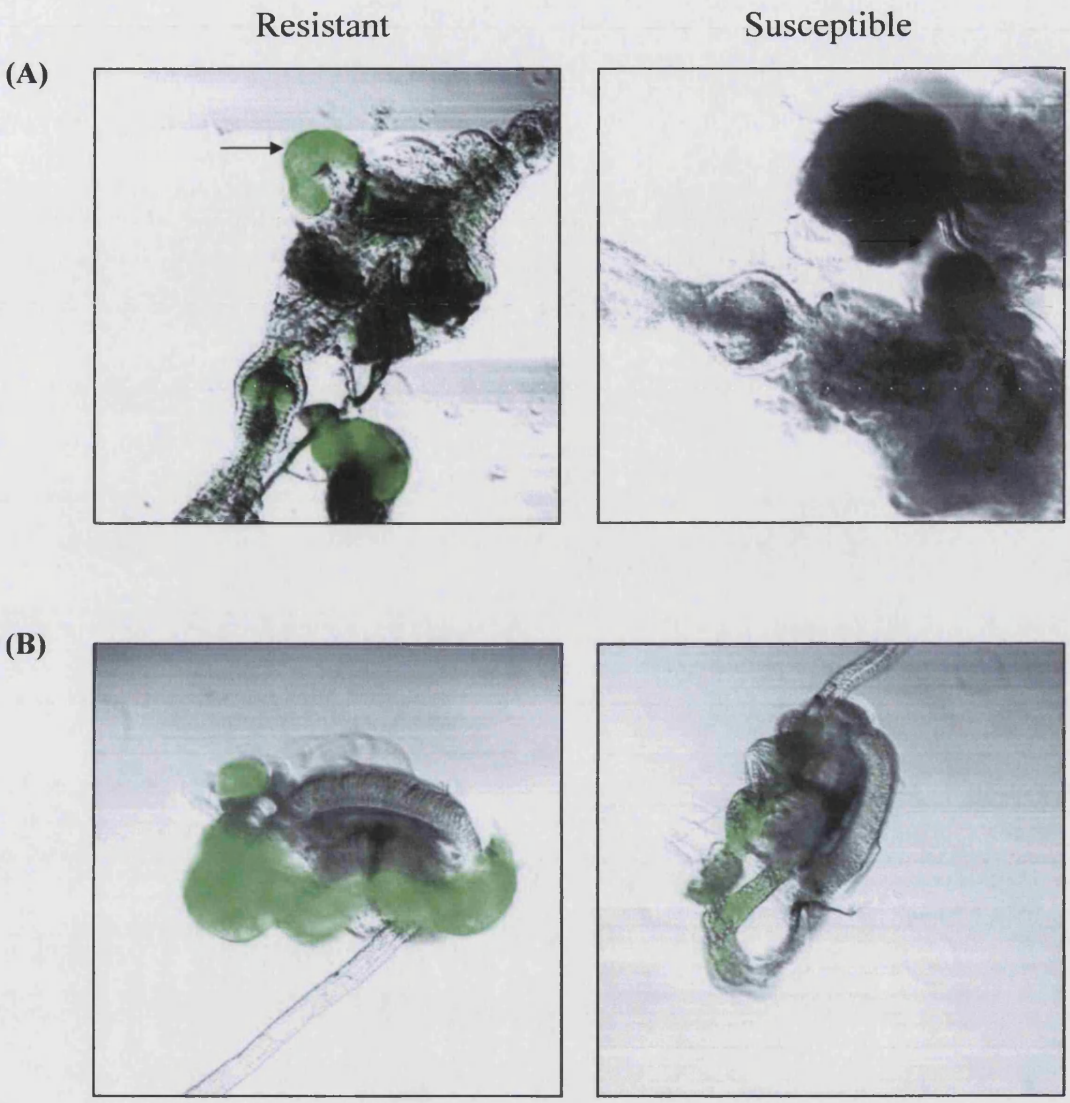


Fig. 5.14 GFP expression driven by *Cyp6g1* promoter regions. **(A)** Confocal microscopy images of GFP expression driven by a GAL4 UAS system using driver strains with 5' Hikone-R (resistant) and 5' Canton-S (susceptible) promoter regions. Images show differential expression of GFP in reproductive tissues from resistant and susceptible constructs. **(A)** Female reproductive tissue with spermathecae indicted (arrows). **(B)** Male reproductive tissues. (Images kindly provided by C. McCart).

Further investigation is required to understand these observations; life-history traits of virgin flies have been studied but will undoubtedly be complicated by the additional factor of male genotype when mated flies are used. To associate the fitness costs and benefits of CYP6G1 overexpression, determining the full range of functions of the protein is of primary importance.

5.4.12 Role of CYP6G1 in Metabolism of Endogenous Compounds

The possibility exists that CYP6G1 contributes to hormone metabolism in *D. melanogaster*. Firstly, a number of insect P450 monooxygenases have previously been implicated in hormone synthesis (e.g. Helvig *et al.*, 2004). Secondly, application of juvenile hormone III to flies elicits a differential response in fecundity between DDT-resistant and susceptible flies (McCart, 2006). Interference with hormone titres would be consistent with the multiple effects on life-history traits observed in Hikone-R. Juvenile hormone has major pleiotropic effects on the *Drosophila* phenotype, and its influences are broad ranging, including ovarian development and fecundity, diapause, metamorphosis, maturation and lifespan (Flatt *et al.*, 2005). An interesting example of the effect of JH on development comes from *Drosophila mercatorum*, in which some naturally occurring genotypes known as aa (abnormal abdomen) have atypical JH elevations due to decreased activity of juvenile hormone esterase, an enzyme that inactivates the hormone (Finch & Rose, 1995). For female flies, the presumed increase in JH titre results in early sexual maturation, precocious fecundity and reduced longevity.

Like CYP6G1, CYP6A1 of the housefly *Musca domestica* is constitutively overexpressed in the midgut and fat body of some insecticide-resistant strains, and has broad substrate specificity (Andersen *et al.*, 1997). Owing to its unusually available haem group, CYP6A1 can metabolise 6,7- and 10,11-epoxides of JH, the two juvenile hormones of *M. domestica* and higher diptera, as well as a variety of pesticides. An imidazole compound that inhibits CYP15A1-mediated JH III synthesis in the corpora allata of cockroaches also inhibits CYP6A1 activity.

In male *Aedes aegypti*, JH III is synthesised by a cytochrome P450 in the male accessory glands, allowing the mosquito to synthesise and secrete the hormone independently of the corpus allatum (Borovsky *et al.*, 1994). This may be of relevance to the role of CYP6G1, which is also expressed in the male accessory glands. Furthermore, the hormones synthesised by the mosquito accessory glands are transferred with the sperm into the female, so modifying development of the ovaries.

As mentioned above, it would be of interest to examine the effect of male genotype on the life-history traits of DDT-resistant and susceptible flies, i.e. to explore the possibility that overexpression of CYP6G1 in the sperm has an effect on, for instance, developmental rates of fertilised embryos. Further analyses with JH III exposure are required, as are measurements of Hikone-R and Canton-S life spans; high titres of juvenile hormone are known to result in shorter life spans in *Drosophila* (Flatt *et al.*, 2005). In *D. melanogaster*, JH levels rise just prior to pupal eclosion, and rapidly reach a peak before the adult matures (Bownes & Rembold, 1987). Sexual maturity is reached, the flies mate and begin laying eggs within two days. By contrast, *Drosophila hydei* have a longer lifecycle and maturation time, and the increase in JH occurs over 10 days. If CYP6G1 does contribute to the synthesis of JH, one can envisage that the high level of transcription of *Cyp6g1* in Hikone-R flies may be associated with increased titres of the hormone and result in the accelerated development of these flies reported by McCart *et al.* (2005) in comparison to Canton-S flies. Consistently, it would be expected that the exogenous supply of JH would bring the life-history traits of Canton-S in line with those of untreated Hikone-R.

The ability of CYP6G1 to metabolise insect hormones has not been fully explored. When juvenile hormone III and ecdysone were docked into our homology model, the former fitted comfortably into the active site but ecdysone did not. A parameter of particular interest in the docking of these compounds is the total complex interaction energy, calculated from the minimised ligand-enzyme complex. Ecdysone has a molecular weight of 464.6Da, and is larger than the insecticides tested. The total complex energy score for ecdysone docked in CYP6G1 was 484.1kcal/mol, which is considerably greater than for the other docked compounds. For example, by comparison DDT and nitenpyram, which are 354.5Da and 270.7Da in molecular weight and are known to be metabolised by CYP6G1, had equivalent complex

energies of –27.7 and 35.4kcal/mol, respectively. Interestingly, juvenile hormone III had an intermediate complex energy score, 23.2kcal/mol, indicating that the hormone fits in the active site and may form favourable interactions. It may be of interest to dock other substrates in the juvenile hormone synthesis pathways, such as methyl farnesoate, which is immediately upstream of JH III in the mevalonate pathway. Given their large size, it is unlikely that intermediates between cholesterol and the principle moulting hormone 20-hydroxyecdysone would fit in the small active site cavity of CYP6G1. It is worth stressing however that cytochrome P450 genes have previously been implicated in both of these pathways (Helvig *et al.*, 2004, Rewitz *et al.*, 2006).

Docking endogenous substrates into the CYP6G1 homology model is a useful first step in understanding the possible activities of the protein outside of insecticide metabolism. It may also be possible to develop a complete reconstituted system to analyse metabolic substrates and products *in vitro*. Lastly, a look at the insect CYP6 family evolution should give insights into its ancestry and possible role in *D. melanogaster* before the introduction of insecticides.

CHAPTER 6

General Discussion and Future Perspectives

6.1 GENERAL DISCUSSION

The predictive quality of homology models is limited. It has been suggested by Scott *et al.* (2001) that the level of structural knowledge necessary to predict the structure-function relationships in P450s may require the atomic structure of a representative from each subfamily. To date, no insect P450 structures have been deposited in the protein data bank, and knowledge of the active site cavities and roles of individual amino acid residues comes only from comparison with bacterial and mammalian P450s, homology modelling and site directed mutagenesis studies.

Our homology models likely show high fidelity in the highly conserved regions, particularly with respect to the placement of secondary structural elements such as helices I and K, while the less conserved and more flexible regions are expected to be less accurate. Crucially, interactions of the active site and the associated complex energy calculations generally show consistency with biological data; i.e., in cases where a specific P450 is able to confer a degree of resistance to a particular insecticide, we found the compound both fitted in the active site and formed favourable interactions. CYP6G1 with docked DDT forms a 'V' shaped cavity that fits well around the bound substrate. The binding mode is characterised by multiple hydrophobic contacts, enabled by the presence of several phenylalanine residues in the cavity, and has been observed previously for proteins able to interact with DDT and other organic polychlorinated compounds (D'Ursi *et al.*, 2005). A proposed hydrophobic access channel is also consistent with expectations for an enzyme able to metabolise such highly hydrophobic substrates. Since DDT is the largest of the substrates we docked, the ability of CYP6G1 to accommodate other insecticides requires only moderate flexibility and the employment of different residues within the cavity.

An important next step with the CYP6G1 homology model presented here is to dock compounds known not to be metabolised by the enzyme, to provide a reference frame against which favourable and unfavourable docking energies can be compared. CYP6A2 was shown by Daborn *et al.* (2007) not to provide resistance to any of the insecticides tested, but for CYP6G1 the relative resistance to diazinon was not determined. It is unclear whether this indicates that the substrate is not metabolised; given the relatively low docking and interaction energies calculated from our model, one might predict that overexpression of CYP6G1 will not confer resistance to this compound. Extending these studies, it will also be interesting to dock other insecticides that the enzyme has been shown to confer resistance to, such as lufenuron and imidacloprid (Daborn *et al.*, 2002b, McCart, 2006). Lastly, insecticides should be docked for which the metabolic activity has not been measured, to test the predictive qualities of the model.

6.2 FUTURE PERSPECTIVES

A structural model of CYP6G1 provides not only an explanation for the ability of the enzyme, through overexpression, to confer resistance to a broad range of insecticides. Our model, and other models of cytochromes P450, particularly those of important pest species, may be used in future studies for the design of novel insecticides. Developing P450-specific inhibitors or compounds that exert similar effects, such as on the ion channels of the insect nervous system, but do not fit into the P450 cavity, is a task for the future. P450s of special interest for such studies include CYP6B7, which is associated with pyrethroid resistance in *Helicoverpa armigera* (Ranasinghe *et al.*, 1998), a lepidopteran responsible for significant crop losses worldwide, and CYP6Z1, implicated in DDT and permethrin resistance in *Anopheles gambiae* (David *et al.*, 2005), a mosquito vector of malaria. The establishment of a reconstituted cytochrome P450 *in vitro* system, as has been used by Jacobsen *et al.*, (2006), will allow verification of the findings.

To further explore the possibility that CYP6G1 contributes to the metabolism of endogenous compounds, a range of hormone intermediary compounds should be docked into our model. *In vivo*, further studies with the exogenous supply of juvenile hormone are needed, as outlined in Chapter 5. It may also be fruitful to measure the hormone titres of Canton-S and Hikone-R flies using gas chromatography-mass spectrometry, as has recently been applied to the study of termite populations (Park & Raina, 2004). Lastly, it would be of interest to investigate CYP6G1 from an evolutionary perspective, with the aim of identifying a role of the protein prior to the introduction of DDT.

CHAPTER 7

References

7.1 REFERENCES

Akhurst, R. J. (1980) Morphological and functional dimorphism in *Xenorhabdus* spp., bacteria symbiotically associated with the insect pathogenic nematodes *Neoplectana* and *Heterorhabditis*. *Journal of General Microbiology* 121: 303-309

Akhurst, R. J. and N. E. Boemare (1988) A numerical taxonomic study of the genus *Xenorhabdus* (*Enterobacteriaceae*) and proposed elevation of the subspecies of *X. nematophilus* to species. *Journal of General Microbiology* 134(7): 751-761

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25(17): 3389-3402

Amano, A., T. Nakamura, S. Kimura, I. Morisaki, I. Nakagawa, S. Kawabata and S. Hamada (1999) Molecular interactions of *Porphyromonas gingivalis* fimbriae with host proteins: Kinetic analyses based on surface plasmon resonance. *Infection and Immunity* 67(5): 2399-2405

Amichot, M., S. Tares, A. Brun-Barale, L. Arthaud, J. M. Bride and J. B. Berge (2004) Point mutations associated with insecticide resistance in the *Drosophila* cytochrome P450 *Cyp6a2* enable DDT metabolism. *Eur J Biochem* 271(7): 1250-1257

Andersen, J. F., J. K. Walding, P. H. Evans, W. S. Bowers and R. Feyereisen (1997) Substrate specificity for the epoxidation of terpenoids and active site topology of house fly cytochrome P450 6A1. *Chemical Research in Toxicology* 10(2): 156-164

Anyanful, A., J. M. Dolan-Livengood, T. Lewis, S. Sheth, M. N. Dezalia, M. A. Sherman, L. V. Kalman, G. M. Benian and D. Kalman (2005) Paralysis and killing of *Caenorhabditis elegans* by enteropathogenic *Escherichia coli* requires the bacterial tryptophanase gene. *Molecular Microbiology* 57(4): 988-1007

Au, C., P. Dean, S. E. Reynolds and R. H. ffrench-Constant (2004) Effect of the insect pathogenic bacterium *Photorhabdus* on insect phagocytes. *Cellular Microbiology* 6(1): 89-95

Bae, E. and G. N. Phillips, Jr. (2004) Structures and analysis of highly homologous psychrophilic, mesophilic, and thermophilic adenylate kinases. *Journal of Biological Chemistry* 279(27): 28202-28208

Bakker, D. P., B. R. Postmus, H. J. Busscher and H. C. van der Mei (2004) Bacterial strains isolated from different niches can exhibit different patterns of adhesion to substrata. *Applied and Environmental Microbiology* 70(6): 3758-3760

Basus, V. (1989) Proton nuclear magnetic resonance assignments. Nuclear Magnetic Resonance. N. J. Oppenheimer and T. L. James. San Diego, Academic Press. 177: 132-147

Baudry, J., W. Li, L. Pan, M. R. Berenbaum and M. A. Schuler (2003) Molecular docking of substrates and inhibitors in the catalytic site of CYP6B1, an insect cytochrome p450 monooxygenase. *Protein Engineering* 16(8): 577-587

Beauchamp, J. C. and N. W. Isaacs (1999) Methods for X-ray diffraction analysis of macromolecular structures. *Current Opinion in Chemical Biology* 3(5): 525-529

Benz, I. and M. A. Schmidt (2002) Never say never again: protein glycosylation in pathogenic bacteria. *Molecular Microbiology* 45(2): 267-276

Bian, Z. and S. Normark (1997) Nucleator function of CsgB for the assembly of adhesive surface organelles in *Escherichia coli*. *EMBO Journal* 16(19): 5827-5836

Bintrim, S. B. and J. C. Ensign (1998) Insertional inactivation of genes encoding the crystalline inclusion proteins of *Photorhabdus luminescens* results in mutants with pleiotropic phenotypes. *Journal of Bacteriology* 180(5): 1261-1269

Blackburn, M., E. Golubeva, D. Bowen and R. H. ffrench-Constant (1998) A novel insecticidal toxin from *Photorhabdus luminescens*, toxin complex a (Tca), and its histopathological effects on the midgut of *Manduca sexta*. *Applied and Environmental Microbiology* 64(8): 3036-3041

Blackburn, M. B., J. M. Domek, D. B. Gelman and J. S. Hu (2005) The broadly insecticidal *Photorhabdus luminescens* toxin complex a (Tca): activity against the Colorado potato beetle, *Leptinotarsa decemlineata*, and sweet potato whitefly, *Bemisia tabaci*. *Journal of Insect Science* 5: 32

Bloomfield, A. A. (1985) Chapter 10, Biological Applications. Dynamic Light Scattering. R. Pecora. New York, Plenum Press: 363-416

Bloomquist, J. R. (1996) Ion channels as targets for insecticides. *Annual Review of Entomology* 41: 163-190

Boemare, N. E. and R. J. Akhurst (1988) Biochemical and physiological characterisation of colony form variants in *Xenorhabdus* spp. (*Enterbacteriaceae*). *Journal of General Microbiology* 134: 1835-1845

Boemare, N. E., R. J. Akhurst and R. G. Mourant (1993) DNA relatedness between *Xenorhabdus* spp. (*Enterobacteriaceae*), symbiotic bacteria of entomopathogenic nematodes, and a proposal to transfer *Xenorhabdus luminescens* to a new genus, *Photorhabdus* gen. nov. *International Journal of Systematic Bacteriology* 43(2): 249-255

Boles, B. R., M. Thoendel and P. K. Singh (2005) Rhamnolipids mediate detachment of *Pseudomonas aeruginosa* from biofilms. *Molecular Microbiology* 57(5): 1210-1223

Bonhivers, M., M. Desmadril, G. S. Moeck, P. Boulanger, A. Colomer-Pallas and L. Letellier (2001) Stability studies of FhuA, a two-domain outer membrane protein from *Escherichia coli*. *Biochemistry* 40(8): 2606-2613

Borovsky, D., D. A. Carlson, R. G. Hancock, H. Rembold and E. van Handel (1994) *De novo* biosynthesis of juvenile hormone III and I by the accessory glands of the male mosquito. *Insect Biochemistry and Molecular Biology* 24(5): 437-444

Bowen, D., T. A. Rocheleau, M. Blackburn, O. Andreev, E. Golubeva, R. Bhartia and R. H. ffrench-Constant (1998) Insecticidal toxins from the bacterium *Photorhabdus luminescens*. *Science* 280(5372): 2129-2132

Bowen, D. J. and J. C. Ensign (1998) Purification and characterization of a high-molecular-weight insecticidal protein complex produced by the entomopathogenic bacterium *Photorhabdus luminescens*. *Applied Environmental Microbiology* 64(8): 3029-3035

Bowen, D. J., T. A. Rocheleau, C. K. Grutzmacher, L. Meslet, M. Valens, D. Marble, A. J. Dowling, R. H. ffrench-Constant and M. A. Blight (2003) Genetic and biochemical characterization of PrtA, an RTX-like metalloprotease from *Photorhabdus*. *Microbiology* 149: 1581-1591

Bownes, M. and H. Rembold (1987) The titre of juvenile hormone during the pupal and adult stages of the life cycle of *Drosophila melanogaster*. *European Journal of Biochemistry* 164(3): 709-712

Bradley, D. E. (1980) Function of *Pseudomonas Aeruginosa* Pao polar pili - twitching motility. *Canadian Journal of Microbiology* 26(2): 146-154

Branda, S. S., Å. Vik, L. Friedman and R. Kolter (2005) Biofilms: the matrix revisited. *Trends in Microbiology* 13(1): 20-26

Branden, C. and J. Tooze (1999) Chapter 18, Determination of protein structure. Introduction to protein structure, second edition. C. Brandon and J. Tooze. New York, Garland publishing: 373-392

Brandt, A., M. Scharf, J. H. F. Pedra, G. Holmes, A. Dean, M. Kreitman and B. R. Pittendrigh (2002) Differential expression and induction of two *Drosophila* cytochrome P450 genes near the Rst(2)DDT locus. *Insect Molecular Biology* 11(4): 337-341

Brombacher, E., A. Baratto, C. Dorel and P. Landini (2006) Gene expression regulation by the Curli activator CsgD protein: modulation of cellulose biosynthesis and control of negative determinants for microbial adhesion. *Journal of Bacteriology* 188(6): 2027-2037

Brugirard-Ricaud, K., E. Duchaud, A. Givaudan, P. A. Girard, F. Kunst, N. Boemare, M. Brehelin and R. Zumbihl (2005) Site-specific antiphagocytic function of the *Photorhabdus luminescens* type III secretion system during insect colonization. *Cellular Microbiology* 7(3): 363-371

Brunham, R. C., F. A. Plummer and R. S. Stephens (1993) Bacterial antigenic variation, host immune-response, and pathogen-host coevolution. *Infection and Immunity* 61(6): 2273-2276

Busscher, H. J. and H. C. van der Mei (1997) Physico-chemical interactions in initial microbial adhesion and relevance for biofilm formation. *Advances in Dental Research* 11(1): 24-32

Cerca, N., G. B. Pier, M. Vilanova, R. Oliveira and J. Azeredo (2005) Quantitative analysis of adhesion and biofilm formation on hydrophilic and hydrophobic surfaces of clinical isolates of *Staphylococcus epidermidis*. *Research in Microbiology* 156(4): 506-514

Chakrabartty, A. and R. L. Baldwin (1995) Stability of alpha-helices. *Advances in Protein Chemistry* 46: 141-176

Chapman, M. R., L. S. Robinson, J. S. Pinkner, R. Roth, J. Heuser, M. Hammar, S. Normark and S. J. Hultgren (2002) Role of *Escherichia coli* curli operons in directing amyloid fiber formation. *Science* 295(5556): 851-855

Chayen, N. E. (1998) Comparative studies of protein crystallization by vapour-diffusion and microbatch techniques. *Acta Crystallographica Section D* 54(1): 8-15

Cheng, J., A. Z. Randall, M. J. Sweredoski and P. Baldi (2005) SCRATCH: a protein structure and structural feature prediction server. *Nucleic Acids Research* 33(Web server issue): W72-W76

Chiti, F., M. Stefani, N. Taddei, G. Ramponi and C. M. Dobson (2003) Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature* 424(6950): 805-808

Chung, H., M. R. Bogwitz, C. McCart, A. Andrianopoulos, R. H. ffrench-Constant, P. Batterham and P. J. Daborn (2007) Cis-regulatory elements in the Accord retrotransposon result in tissue-specific expression of the *Drosophila melanogaster* insecticide resistance gene *Cyp6g1*. *Genetics* 175(3): 1071-1077

Ciche, T. A. and J. C. Ensign (2003) For the insect pathogen *Photorhabdus luminescens*, which end of a nematode is out? *Applied and Environmental Microbiology* 69(4): 1890-1897

Ciche, T. A., S. B. Bintrim, A. R. Horswill and J. C. Ensign (2001) A phosphopantetheinyl transferase homolog is essential for *Photorhabdus luminescens* to support growth and reproduction of the entomopathogenic nematode *Heterorhabditis bacteriophora*. *Journal of Bacteriology* 183(10): 3117-3126

Clarke, D. J. and B. C. A. Dowds (1995) Virulence mechanisms of *Photorhabdus* sp strain K122 toward wax moth larvae. *Journal of Invertebrate Pathology* 66(2): 149-155

Conway, K. A., S. J. Lee, J. C. Rochet, T. T. Ding, R. E. Williamson and P. T. Lansbury, Jr. (2000) Acceleration of oligomerization, not fibrillization, is a shared property of both alpha-synuclein mutations linked to early-onset Parkinson's disease: implications for pathogenesis and therapy. *Proceedings of the National Academy of Sciences* 97(2): 571-576

Cooper, A. (1999) Thermodynamic analysis of biomolecular interactions. *Current Opinion in Chemical Biology* 3: 557-563

Cooper, A. (2001) DSC interpretation notes. University of Glasgow, Department fo Chemistry. Glasgow: 1-4

Cosme, J. and E. F. Johnson (2000) Engineering microsomal cytochrome P450 2C5 to be a soluble, monomeric enzyme. Mutations that alter aggregation, phospholipid dependence of catalysis, and membrane binding. *Journal of Biological Chemistry* 275(4): 2545-2553

Cosme, J., A. Ward, L. Vuillard, P. Williams and B. Hamilton (2003) Methods for purifcation of cytochrome P450 proteins and their crystallisation. Patent Corporation Treaty (PCT/GB02/02668)

Costerton, J. W., P. S. Stewart and E. P. Greenberg (1999) Bacterial biofilms: A common cause of persistent infections. *Science* 284(5418): 1318-1322

Couillault, C. and J. J. Ewbank (2002) Diverse bacteria are pathogens of *Caenorhabditis elegans*. *Infection and Immunity* 70(8): 4705-4707

Coustau, C., C. Chevillon and R. ffrench-Constant (2000) Resistance to xenobiotics and parasites: can we count the cost? *Trends in Ecology and Evolution* 15(9): 378-383

Creagh, A. L., J. Koska, P. E. Johnson, P. Tomme, M. D. Joshi, L. P. McIntosh, D. G. Kilburn and C. A. Haynes (1998) Stability and oligosaccharide binding of the N1 cellulose-binding domain of *Cellulomonas fimi* endoglucanase CenC. *Biochemistry* 37(10): 3529-3537

Creighton, T. E. (1993a) Chapter 6, The folded conformation of globular proteins. Proteins, Structures and molecular properties, Second edition. T. E. Creighton. New York, W.H. Freeman and Company: 201-260

Creighton, T. E. (1993b) Chapter 7, Proteins in solution and in membranes. Proteins, Structures and molecular properties, Second edition. T. E. Creighton. New York, W.H. Freeman and Company: 201-260

Daborn, P. J., C. Lumb, A. Boey, W. Wong, R. H. ffrench-Constant and P. Batterham (2007) Evaluating the insecticide resistance potential of eight *Drosophila melanogaster* cytochrome P450 genes by transgenic over-expression. *Insect Biochemistry and Molecular Biology* 37(5): 512-519

Daborn, P. J., N. R. Waterfield, C. P. Silva, C. P. Y. Au, S. Sharma and R. H. ffrench-Constant (2002a) A single *Photorhabdus* gene, makes caterpillars floppy (mcf), allows *Escherichia coli* to persist within and kill insects. *Proceedings of the National Academy of Sciences* 99(16): 10742-10747

Daborn, P. J., J. L. Yen, M. R. Bogwitz, G. Le Goff, E. Feil, S. Jeffers, N. Tijet, T. Perry, D. Heckel, P. Batterham, R. Feyereisen, T. G. Wilson and R. H. ffrench-Constant (2002b) A single p450 allele associated with insecticide resistance in *Drosophila. Science* 297(5590): 2253-2256

Darboe, N., R. Kenjale, W. L. Picking, W. D. Picking and C. R. Middaugh (2006) Physical characterization of MxiH and PrgI, the needle component of the type III secretion apparatus from *Shigella* and *Salmonella. Protein Science* 15(3): 543-552

Danese, P. N., L. A. Pratt, S. L. Dove and R. Kolter (2000) The outer membrane protein, antigen 43, mediates cell-to-cell interactions within *Escherichia coli* biofilms. *Molecular Microbiology* 37(2): 424-432

Darby, C., J. W. Hsu, N. Ghori and S. Falkow (2002) *Caenorhabditis elegans* - Plague bacteria biofilm blocks food intake. *Nature* 417(6886): 243-244

Darby, C., C. L. Cosma, J. H. Thomas and C. Manoil (1999) Lethal paralysis of *Caenorhabditis elegans* by *Pseudomonas aeruginosa. Proceedings of the National Academy of Sciences* 96(26): 15202-15207

D'Arcy, A., C. Elmore, M. Stihle and J. E. Johnston (1996) A novel approach to crystallising proteins under oil. *Journal of Crystal Growth* 168: 175-180

Davail, S., G. Feller, E. Narinx and C. Gerday (1994) Cold adaptation of proteins. Purification, characterization, and sequence of the heat-labile subtilisin from the antarctic psychrophile *Bacillus* TA41. *Journal of Biological Chemistry* 269(26): 17448-17453

David, J. P., C. Strode, J. Vontas, D. Nikou, A. Vaughan, P. M. Pignatelli, C. Louis, J. Hemingway and H. Ranson (2005) The *Anopheles gambiae* detoxification chip: a highly specific microarray to study metabolic-based insecticide resistance in malaria vectors. *Proceedings of the National Academy of Sciences* 102(11): 4080-4084

Davies, D. G., M. R. Parsek, J. P. Pearson, B. H. Iglewski, J. W. Costerton and E. P. Greenberg (1998) The involvement of cell-to-cell signals in the development of a bacterial biofilm. *Science* 280(5361): 295-298

de Vocht, M. L., I. Reviakine, H. A. Wosten, A. Brisson, J. G. Wessels and G. T. Robillard (2000) Structural and functional role of the disulfide bridges in the hydrophobin SC3. *Journal of Biological Chemistry* 275(37): 28428-28432

Deane, J. E., P. Roversi, F. S. Cordes, S. Johnson, R. Kenjale, S. Daniell, F. Booy, W. D. Picking, W. L. Picking, A. J. Blocker and S. M. Lea (2006) Molecular model of a type III secretion system needle: Implications for host-cell sensing. *Proceedings of the National Academy of Sciences* 103(33): 12529-12533

Delcarte, J., M. Fauconnier, P. Jacques, K. Matsui, P. Thonart and M. Marlier (2003) Optimisation of expression and immobilized metal ion affinity chromatographic purification of recombinant (His)6-tagged cytochrome P450 hydroperoxide lyase in *Escherichia coli*. *Journal of Chromatography B* 786(1-2): 229-236

Doerr, A. (2006) Together at last: crystallography and NMR. *Nature Methods* 3(2): 6

Doig, A. J. and M. J. E. Sternberg (1995) Side-chain conformational entropy in protein folding. *Protein Science* 4(11): 2247-2251

Doig, P., T. Todd, P. A. Sastry, K. K. Lee, R. S. Hodges, W. Paranchych and R. T. Irvin (1988) Role of pili in adhesion of *Pseudomonas aeruginosa* to human respiratory epithelial cells. *Infection and Immunity* 56(6): 1641-1646

Doray (2001) N-terminal deletions and His-tag fusions dramatically affect expression of cytochrome p450 2C2 in bacteria. *Archives of Biochemistry and Biophysics* 393(1): 143-153

Dowling, A. J., P. J. Daborn, N. R. Waterfield, P. Wang, C. H. Streuli and R. H. ffrench-Constant (2004) The insecticidal toxin Makes caterpillars floppy (Mcf) promotes apoptosis in mammalian cells. *Cellular Microbiology* 6(4): 345-353

Duchaud, E., C. Rusniok, L. Frangeul, C. Buchrieser, A. Givaudan, S. Taourit, S. Bocs, C. Boursaux-Eude, M. Chandler, J. F. Charles, E. Dassa, R. Derose, S. Derzelle, G. Freyssinet, S. Gaudriault, C. Medigue, A. Lanois, K. Powell, P. Siguier, R. Vincent, V. Wingate, M. Zouine, P. Glaser, N. Boemare, A. Danchin and F. Kunst (2003) The genome sequence of the entomopathogenic bacterium *Photorhabdus luminescens*. *Nature Biotechnology* 21(11): 1307-1313

Dunne, W. M., Jr. (2002) Bacterial adhesion: Seen any good biofilms lately? *Clinical Microbiology Reviews* 15(2): 155-166

D'Ursi, P., E. Salvi, P. Fossa, L. Milanesi and E. Rovida (2005) Modelling the interaction of steroid receptors with endocrine disrupting chemicals. *BMC Bioinformatics* 6 Suppl 4: S10

Dybvig, K. (1993) DNA rearrangements and phenotypic switching in prokaryotes. *Molecular Microbiology* 10(3): 465-471

Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32(5): 1792-1797

Edwards, R. A. and J. L. Puente (1998) Fimbrial expression in enteric bacteria: a critical step in intestinal pathogenesis. *Trends in Microbiology* 6(7): 282-287

Eldefrawi, M. E., R. Miskus and V. Sutcher (1960) Methylenedioxyphenyl derivatives as synergists for carbamate insecticides on susceptible, DDT-and Parathion-Resistant House Flies. *Journal of Economic Entomology* 53: 231-234

Eleftherianos, I., J. Marokhazi, P. J. Millichap, A. J. Hodgkinson, A. Sriboonlert, R. H. ffrench-Constant and S. E. Reynolds (2006) Prior infection of *Manduca sexta* with non-pathogenic *Escherichia coli* elicits immunity to pathogenic *Photorhabdus luminescens*: Roles of immune-related proteins shown by RNA interference. *Insect Biochemistry and Molecular Biology* 36(6): 517-525

Eleftherianos, I., S. Boundy, S. A. Joyce, S. Aslam, J. W. Marshall, R. J. Cox, T. J. Simpson, D. J. Clarke, R. H. ffrench-Constant and S. E. Reynolds (2007) An antibiotic produced by an insect-pathogenic bacterium suppresses host defenses through phenoloxidase inhibition. *Proceedings of the National Academy of Sciences* (7): 2419-2424

Ellis, J. (1987) Proteins as molecular chaperones. *Nature* 328(6129): 378-379

Ellis, R. J. (2006) Protein folding: inside the cage. *Nature* 442(7101): 360-362

Encinar, J. A., A. Fernandez, J. A. Ferragut, J. M. Gonzalez-Ros, B. R. DasGupta, M. Montal and A. Ferrer-Montiel (1998) Structural stabilization of botulinum neurotoxins by tyrosine phosphorylation. *FEBS Letters* 429(1): 78-82

Fan, J. and L. Nan (2007) Protein structural codes and nucleation sites for protein folding. *Chinese Physics* 16(2): 392-404

Fändrich, M. (2007) Absolute correlation between lag time and growth rate in the spontaneous formation of several amyloid-like aggregates and fibrils. *Journal of Molecular Biology* 365(5): 1266-1270

Farmer, J. J., J. H. Jorgensen, P. A. D. Grimont, R. J. Akhurst, G. O. Poinar, E. Ageron, G. V. Pierce, J. A. Smith, G. P. Carter, K. L. Wilson and F. W. Hickmanbrenner (1989) *Xenorhabdus luminescens* (DNA Hybridization Group-5) from Human Clinical Specimens. *Journal of Clinical Microbiology* 27(7): 1594-1600

Feller, G., P. Sonnet and C. Gerday (1995) The beta-lactamase secreted by the antarctic psychrophile *Psychrobacter immobilis* A8. *Applied and Environmental Microbiology* 61(12): 4474-4476

Fernandez-Escamilla, A. M., F. Rousseau, J. Schymkowitz and L. Serrano (2004) Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nature Biotechnology* 22(10): 1302-1306

Ferré-D'Amaré, A. R. and S. K. Burley (1997) Dynamic light scattering in evaluating crystallizability of macromolecules. *Methods in Enzymology* 276: 157-166

Ferrieres, L. and D. J. Clarke (2003) The RcsC sensor kinase is required for normal biofilm formation in *Escherichia coli* K-12 and controls the expression of a regulon in response to growth on a solid surface. *Molecular Microbiology* 50(5): 1665-1682

Feyereisen, R. (1999) Insect P450 enzymes. *Annual Review of Entomology* 44: 507-533

ffrench-Constant, R. H., P. J. Daborn and G. Le Goff (2004) The genetics and genomics of insecticide resistance. *Trends in Genetics* 20(3): 163-170

ffrench-Constant, R. H., B. Pittendrigh, A. Vaughan and N. Anthony (1998) Why are there so few resistance-associated mutations in insecticide target genes? *Philosophical Transactions of the Royal Society of London Series B* 353: 1685-1693

ffrench-Constant, R. H., T. A. Rocheleau, J. C. Steichen and A. E. Chalmers (1993) A point mutation in a Drosophila GABA receptor confers insecticide resistance. *Nature* 363(6428): 449-51

ffrench-Constant, R. H. and N. R. Waterfield (2006) Ground control for insect pests. *Nature Biotechnology* 24(6): 660-661

Finch, C. E. and M. R. Rose (1995) Hormones and the physiological architecture of life history evolution. *Quarterly Review of Biology* 70(1): 1-52

Finkel, S. E. (2006) Long-term survival during stationary phase: evolution and the GASP phenotype. *Nature Reviews Microbiology* 4(2): 113-120

Finn, R. D., J. Mistry, B. Schuster-Bockler, S. Griffiths-Jones, V. Hollich, T. Lassmann, S. Moxon, M. Marshall, A. Khanna, R. Durbin, S. R. Eddy, E. L. L. Sonnhammer and A. Bateman (2006) Pfam: clans, web tools and services. *Nucleic Acids Research* 34(supplement 1): D247-251

Fischer-Le Saux, M., V. Viallard, B. Brunel, P. Normand and N. E. Boemare (1999) Polyphasic classification of the genus *Photorhabdus* and proposal of new taxa: *P. luminescens* subsp *luminescens* subsp nov., *P. luminescens* subsp *akhurstii* subsp nov., *P. luminescens* subsp *laumondii* subsp nov., *P. temperata* sp nov., *P. temperata* subsp *temperata* subsp nov and *P. asymbiotica* sp nov. *International Journal of Systematic Bacteriology* 49: 1645-1656

Flatt, T., M. P. Tu and M. Tatar (2005) Hormonal pleiotropy and the juvenile hormone regulation of *Drosophila* development and life history. *Bioessays* 27(10): 999-1010

Ford, N. C. J. (1985) Chapter 2, Light scattering apparatus. Dynamic Light Scattering. R. Pecora. New York, Plenum Press: 7-58

Forst, S., B. Dowds, N. Boemare and E. Stackebrandt (1997) *Xenorhabdus* and *Photorhabdus* spp.: Bugs that kill bugs. *Annual Review of Microbiology* 51: 47-72

Foster, M. P., C. A. McElroy and C. D. Amero (2007) Solution NMR of large molecules and assemblies. *Biochemistry* 46(2): 331-340

Friedman, L. and R. Kolter (2004) Two genetic loci produce distinct carbohydrate-rich structural components of the *Pseudomonas aeruginosa* biofilm matrix. *Journal of Bacteriology* 186(14): 4457-4465

Ganesh, C., A. N. Shah, C. P. Swaminathan, A. Surolia and R. Varadarajan (1997) Thermodynamic characterization of the reversible, two-state unfolding of maltose binding protein, a large two-domain protein. *Biochemistry* 36(16): 5020-5028

Gasteiger, E., C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel and A. Bairoch (2005) Protein Identification and Analysis Tools on the ExPASy Server;. The Proteomics Protocols Handbook. I. M. Walker. Totowa, NJ, Humana Press: 571-607

Gaudriault, S., E. Duchaud, A. Lanois, A. S. Canoy, S. Bourot, R. DeRose, F. Kunst, N. Boemare and A. Givaudan (2006) Whole-genome comparison between *Photorhabdus* strains to identify genomic regions involved in the specificity of nematode interaction. *Journal of Bacteriology* 188(2): 809-814

Gavel, O. Y., S. A. Bursakov, D. G. Pina, G. G. Zhadan, J. J. Moura, I. Moura and V. L. Shnyrov (2004) Structural stability of adenylate kinase from the sulfate-reducing bacteria *Desulfovibrio gigas*. *Biophysical Chemistry* 110(1-2): 83-92

Georgiou, G. and P. Valax (1996) Expression of correctly folded proteins in *Escherichia coli*. *Current Opinion in Biotechnology* 7(2): 190-197

Georlette, D., V. Blaise, C. Dohmen, F. Bouillenne, B. Damien, E. Depiereux, C. Gerday, V. N. Uversky and G. Feller (2003) Cofactor binding modulates the conformational stabilities and unfolding patterns of NAD(+)-dependent DNA ligases from *Escherichia coli* and *Thermus scotoductus*. *Journal of Biological Chemistry* 278(50): 49945-49953

Gerrard, J. G., S. A. Joyce, D. J. Clarke, R. H. ffrench-Constant, G. R. Nimmo, D. F. Looke, E. J. Feil, L. Pearce and N. R. Waterfield (2006) Nematode symbiont for *Photorhabdus asymbiotica*. *Emerging Infectious Diseases* 12(10): 1562-1564

Gerrard, J. G., R. Vohra and G. R. Nimmo (2003) Identification of *Photorhabdus asymbiotica* in cases of human infection. *Communicable Diseases Intelligence* 27(4): 540-541

Glättli, A., X. Daura, D. Seebach and W. F. van Gunsteren (2002) Can one derive the conformational preference of a beta-peptide from its CD spectrum? *Journal of the American Chemical Society* 124(44): 12972-12978

Glieder, A., E. T. Farinas and F. H. Arnold (2002) Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase. *Nature Biotechnology* 20(11): 1135-1139

Gonzalez, A. (2003) Optimizing data collection for structure determination. *Acta Crystallographica Section D* 59(11): 1935-1942

Gotoh, O. (1992) Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. *Journal of Biological Chemistry* 267(1): 83-90

Gravato-Nobre, M. J., H. R. Nicholas, R. Nijland, D. O'Rourke, D. E. Whittington, K. J. Yook and J. Hodgkin (2005) Multiple genes affect sensitivity of *Caenorhabditis elegans* to the bacterial pathogen *Microbacterium nematophilum*. *Genetics* 171(3): 1033-1045

Green, S., A. K. Meeker and D. Shortle (1992) Contributions of the polar, uncharged amino acids to the stability of Staphylococcal nuclease: Evidence for mutational effects on the free energy of the denatured state. *Biochemistry* 31: 5717-5728

Greenfield, N. J. (1996) Methods to estimate the conformation of proteins and polypeptides from circular dichroism data. *Analytical Biochemistry* 235(1): 1-10

Grimont, P. A. D., A. G. Steigerwalt, N. E. Boemare, F. W. Hickman-Brenner, C. Deval and D. J. Brenner (1984) Deoxiribonucleic acid relatedness and phenotypic study of the genus *Xenorhabdus*. *International Journal of Systematic Bacteriology* 34(4): 378-388

Guzman, L. M., D. Belin, M. J. Carson and J. Beckwith (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose pBAD promoter. *Journal of Bacteriology* 177(14): 4121-4130

Hakanpaa, J., A. Paananen, S. Askolin, T. Nakari-Setala, T. Parkkinen, M. Penttila, M. B. Linder and J. Rouvinen (2004) Atomic resolution structure of the HFBII hydrophobin, a self-assembling amphiphile. *Journal of Biological Chemistry* 279(1): 534-539

Hammar, M., Z. Bian and S. Normark (1996) Nucleator-dependent intercellular assembly of adhesive curli organelles in *Escherichia coli*. *Proceedings of the National Academy of Sciences* 93(13): 6562-6566

Han, R. C. and R. U. Ehlers (2001) Effect of *Photorhabdus luminescens* phase variants on the in vivo and in vitro development and reproduction of the entomopathogenic nematodes *Heterorhabditis bacteriophora* and *Steinernema carpocapsae*. *Fems Microbiology Ecology* 35(3): 239-247

Han, R. C., W. Wouts and L. Li (1991) Development and virulence of *Heterorhabditis* spp. strains associated with different *Xenorhabdus luminescens* isolates. *Journal of Invertebrate Pathology* 58(1): 27-32

Hausner, M. and S. Wuertz (1999) High rates of conjugation in bacterial biofilms as determined by quantitative in situ analysis. *Applied and Environmental Microbiology* 65(8): 3710-3713

Helvig, C., J. F. Koener, G. C. Unnithan and R. Feyereisen (2004) CYP15A1, the cytochrome P450 that catalyzes epoxidation of methyl farnesoate to juvenile hormone III in cockroach corpora allata. *Proceedings of the National Academy of Sciences* 101(12): 4024-4029

Hemingway, J., N. J. Hawkes, L. McCarroll and H. Ranson (2004) The molecular basis of insecticide resistance in mosquitoes. *Insect Biochemistry and Molecular Biology* 34(7): 653-665

Henderson, I. R., P. Owen and J. P. Nataro (1999) Molecular switches - the ON and OFF of bacterial phase variation. *Molecular Microbiology* 33(5): 919-932

Hendrickson, W. A., J. R. Horton and D. M. LeMaster (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO Journal* 9(5): 1665-1672

Hennessey, J. P. and W. C. Johnson (1981) Information content in the circular dichroism of proteins. *Biochemistry* 20(5): 1085-1094

Hinnebusch, B. J., R. D. Perry and T. G. Schwan (1996) Role of the *Yersinia pestis* hemin storage (hms) locus in the transmission of plague by fleas. *Science* 273(5273): 367-370

Hober, S., K. Nord and M. Linhult (2007) Protein A chromatography for antibody purification. *Journal of Chromatography B* 848(1): 40-47

Hodgkin, J., P. E. Kuwabara and B. Corneliussen (2000) A novel bacterial pathogen, *Microbacterium nematophilum*, induces morphological change in the nematode *C. elegans*. *Current Biology* 10(24): 1615-1618

Höflich, J., P. Berninsone, C. Gobel, M. J. Gravato-Nobre, B. J. Libby, C. Darby, S. M. Politz, J. Hodgkin, C. B. Hirschberg and R. Baumeister (2004) Loss of srf-3-encoded nucleotide sugar transporter activity in *Caenorhabditis elegans* alters surface antigenicity and prevents bacterial adherence. *Journal of Biological Chemistry* 279(29): 30440-30448

Holmes, K. C. (1999) Structural biology. *Philosophical Transactions of the Royal Society of London Series B* 354(1392): 1977-1984

Honeybee Genome Sequencing Consortium (2006) Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature* 443: 931-949

Hornemann, S., C. Korth, B. Oesch, R. Riek, G. Wider, K. Wuthrich and R. Glockshuber (1997) Recombinant full-length murine prion protein, mPrP(23-231): purification and spectroscopic characterization. *FEBS Letters* 413(2): 277-281

Hunt, S. M., E. M. Werner, B. Huang, M. A. Hamilton and P. S. Stewart (2004) Hypothesis for the Role of Nutrient Starvation in Biofilm Detachment. *Applied and Environmental Microbiology* 70(12): 7418-7425

Hurst, M. R. H., T. R. Glare and T. A. Jackson (2004) Cloning *Serratia entomophila* antifeeding genes - a putative defective prophage active against the grass grub *Costelytra zealandica*. *Journal of Bacteriology* 186(15): 5116-5128

Hurst, M. R. H., T. R. Glare, T. A. Jackson and C. W. Ronson (2000) Plasmid-located pathogenicity determinants of *Serratia entomophila*, the causal agent of amber disease of grass grub, show similarity to the insecticidal toxins of *Photorhabdus luminescens*. *Journal of Bacteriology* 182(18): 5127-5138

Ivanova, N., A. Sorokin, I. Anderson, N. Galleron, B. Candelon, V. Kapatral, A. Bhattacharyya, G. Reznik, N. Mikhailova, A. Lapidus, L. Chu, M. Mazur, E. Goltsman, N. Larsen, M. D'Souza, T. Walunas, Y. Grechkin, G. Pusch, R. Haselkorn, M. Fonstein, S. Dusko Ehrlich, R. Overbeek and N. Kyrpides (2003) Genome sequence of *Bacillus cereus* and comparative analysis with *Bacillus anthracis*. *Nature* 423(6935): 87-91

Jelsch, C., M. M. Teeter, V. Lamzin, V. Pichon-Pesme, R. H. Blessing and C. Lecomte (2000) Accurate protein crystallography at ultra-high resolution: valence electron distribution in crambin. *Proceedings of the National Academy of Sciences* 97(7): 3171-3176

Jenkins, A. T., H. A. Dash, S. Boundy, C. M. Halliwell and R. H. ffrench-Constant (2006) Methoxy-resorufin ether as an electrochemically active biological probe for cytochrome P450 O-demethylation. *Bioelectrochemistry* 68(1): 67-71

Jenkins, A. T., A. Buckling, M. McGhee and R. H. ffrench-Constant (2005) Surface plasmon resonance shows that type IV pili are important in surface attachment by *Pseudomonas aeruginosa*. *Journal of the Royal Society, Interface* 2(3): 255-259

Jenkins, A. T., R. ffrench-constant, A. Buckling, D. J. Clarke and K. Jarvis (2004) Study of the attachment of *Pseudomonas aeruginosa* on gold and modified gold surfaces using surface plasmon resonance. *Biotechnology Progress* 20(4): 1233-1236

Johnson, W. C. J. (1999) Analyzing protein circular dichroism spectra for accurate secondary structures. *Proteins: Structure, Function, and Genetics* 35(3): 307-312

Jones, T. A., J.-Y. Zou, S. W. Cowan and M. Kjeldgaard (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallographica Section A* 47(2): 110-115

Joshua, G. W. P., A. V. Karlyshev, M. P. Smith, K. E. Isherwood, R. W. Titball and B. W. Wren (2003) A *Caenorhabditis elegans* model of *Yersinia* infection: biofilm formation on a biotic surface. *Microbiology* 149: 3221-3229

Joyce, S. A., R. J. Watson and D. J. Clarke (2006) The regulation of pathogenicity and mutualism in *Photorhabdus*. *Current Opinion in Microbiology* 9(2): 127-132

Joyce, S. A. and D. J. Clarke (2003) A hexA homologue from *Photorhabdus* regulates pathogenicity, symbiosis and phenotypic variation. *Molecular Microbiology* 47(5): 1445-1457

Kagawa, N., Q. Cao and K. Kusano (2003) Expression of human aromatase (CYP19) in *Escherichia coli* by N-terminal replacement and induction of cold stress response. *Steroids* 68(2): 205-209

Kapust, R. B. and D. S. Waugh (1999) *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Science* 8(8): 1668-1674

Kasai, S., I. S. Weerashinghe, T. Shono and M. Yamakawa (2000) Molecular cloning, nucleotide sequence and gene expression of a cytochrome P450 (CYP6F1) from the pyrethroid-resistant mosquito, *Culex quinquefasciatus* Say. *Insect Biochemistry and Molecular Biology* 30(2): 163-71

Kasai, S. and J. G. Scott (2000) Overexpression of cytochrome P450 CYP6D1 is associated with monooxygenase-mediated pyrethroid resistance in house flies from Georgia. *Pesticide Biochemistry and Physiology* 68(1): 34-41

Katsikogianni, M. and Y. F. Missirlis (2004) Concise review of mechanisms of bacterial adhesion to biomaterials and of techniques used in estimating bacteria-material interactions. *European Cells and Materials* 8: 37-57

Kelly, S. M., T. J. Jess and N. C. Price (2005) How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1751(2): 119-139

Kelly, S. M. and N. C. Price (2000) The use of circular dichroism in the investigation of protein structure and function. *Current Protein and Peptide Science* 1(4): 349-384

Kendrew, J. C., R. E. Dickerson, B. E. Strandberg, R. G. Hart, D. R. Davies, D. C. Phillips and V. C. Shore (1960) Structure of Myoglobin: A Three-Dimensional Fourier Synthesis at 2 Å. Resolution. 185(4711): 422-427

Kendrew, J. C., G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff and D. C. Phillips (1958) A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* 181(4610): 662-666

Kleywegt, G. J. and T. A. Jones (1994) Detection, delineation, measurement and display of cavities in macromolecular structures. *Acta Crystallographica Section D* 50(2): 178-185

Knoll, W., M. Zizlsperger, T. Liebermann, S. Arnold, A. Badia, M. Liley, D. Piscevic, F. J. Schmitt and J. Spinke (2000) Streptavidin arrays as supramolecular architectures in surface-plasmon optical sensor formats. *Colloids and Surfaces a-Physicochemical and Engineering Aspects* 161(1): 115-137

Koh, K. S., K. W. Lam, M. Alhede, S. Y. Queck, M. Labbate, S. Kjelleberg and S. A. Rice (2007) Phenotypic diversification and adaptation of *Serratia marcescens* MG1 biofilm-derived morphotypes. *Journal of Bacteriology* 189(1): 119-130

Köhler, T., L. K. Curty, F. Barja, C. van Delden and J.-C. Pechere (2000) Swarming of *Pseudomonas aeruginosa* is dependent on cell-to-cell signaling and requires flagella and pili. *Journal of Bacteriology* 182(21): 5990-5996

Koide, S., Z. Bu, D. Risal, T. N. Pham, T. Nakagawa, A. Tamura and D. M. Engelman (1999) Multistep denaturation of *Borrelia burgdorferi* OspA, a protein containing a single-layer beta-sheet. *Biochemistry* 38(15): 4757-4767

Krittanai, C. and W. C. J. Johnson (1997) Correcting the circular dichroism spectra of peptides for contributions of absorbing side chains. *Analytical Biochemistry* 253(1): 57-64

Kuhn, P., M. Knapp, S. M. Soltis, G. Ganshaw, M. Thoene and R. Bott (1998) The 0.78 A structure of a serine protease: Bacillus lentus subtilisin. *Biochemistry* 37(39): 13446-13452

Kumar, S., C.-J. Tsai and R. Nussinov (2000) Factors enhancing protein thermostability. *Protein Engineering* 13(3): 179-191

Kumar, S., C. S. Chen, D. J. Waxman and J.R Halpert (2005) Directed evolution of mammalian cytochrome P450 2B1: mutations outside of the active site enhance the metabolism of several substrates, including the anticancer prodrugs cyclophosphamide and ifosfamide. *Journal of Biological Chemistry* 280 (20): 19569-19575

Kusano, K., M. R. Waterman, M. Sakaguchi, T. Omura and N. Kagawa (1999) Protein Synthesis Inhibitors and Ethanol Selectively Enhance Heterologous Expression of P450s and Related Proteins in *Escherichia coli. Archives of Biochemistry and Biophysics* 367(1): 129-136

Labbate, M., H. Zhu, L. Thung, R. Bandara, M. R. Larsen, M. D. P. Wilcox, M. Givskov, S. A. Rice and S. Kjelleberg (2007) Quorum-sensing regulation of adhesion in *Serratia marcescens* MG1 is surface dependent. *Journal of Bacteriology* 189(7): 2702-2711

Laskowski, R. A., M. W. MacArthur, D. S. Moss and J. M. Thornton (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* 26: 283-291

Le Goff, G., S. Boundy, P. J. Daborn, J. Yen, L. Sofer, R. Lind, C. Sabourault, L. Madi-Ravazzi and R. H. ffrench-Constant (2003) Microarray analysis of cytochrome P450 mediated insecticide resistance in *Drosophila. Insect Biochemistry and Molecular Biology* 33(7): 701-708

Lee, H. J., Y. Yan, G. Marriott and R. M. Corn (2005) Quantitative functional analysis of protein complexes on surfaces. *Journal of Physiology* 563(1): 61-71

Lee, J. and G. Kaletunc (2002) Evaluation of the heat Inactivation of *Escherichia coli* and *Lactobacillus plantarum* by differential scanning calorimetry. *Applied and Environmental Microbiology* 68(11): 5379-5386

Levitt, M. and C. Chothia (1976) Structural patterns in globular proteins. *Nature* 261((5561)): 552-558

Li, H. and X. Zhang (2005) Characterization of thermostable lipase from thermophilic *Geobacillus* sp. TW1. *Protein Expression and Purification* 42(1): 153-159

Li, W. T., R. A. Grayling, K. Sandman, S. Edmondson, J. W. Shriver and J. N. Reeve (1998) Thermodynamic stability of archaeal histones. *Biochemistry* 37(30): 10563-10572

Li, X., C. V. Lockatell, D. E. Johnson and H. L. T. Mobley (2002) Identification of MrpI as the sole recombinase that regulates the phase variation of MR/P fimbria, a bladder colonization factor of uropathogenic *Proteus mirabilis*. *Molecular Microbiology* 45(3): 865-874

Liehl, P., M. Blight, N. Vodovar, F. Boccard and B. Lemaitre (2006) Prevalence of local immune response against oral infection in a *Drosophila/Pseudomonas* infection model. *Plos Pathogens* 2(6): 551-561

Liu, J., R. Berry, G. Poinar and A. Moldenke (1997) Phylogeny of *Photorhabdus* and *Xenohabdus* species and strains as determined by comparison of partial 16S rRNA gene sequences. *International Journal of Systematic Bacteriology* 47(4): 948-951

Lobley, A., L. Whitmore and B. A. Wallace (2002) DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics* 18(1): 211-212

Lu, A., S. B. West, M. Vore, D. Ryan and W. Levin (1974) Role of cytochrome b5 in hydroxylation by a reconstituted cytochrome P-450-containing system. *Journal of Biological Chemistry* 249(21): 6701-6709

Lüdemann, S. K., V. Lounnas and R. C. Wade (2000) How do substrates enter and products exit the buried active site of cytochrome P450cam? 1. Random expulsion molecular dynamics investigation of ligand access channels and mechanisms. *Journal of Molecular Biology* 303(5): 797-811

Luft, J. R., R. J. Collins, N. A. Fehrman, A. M. Lauricella, C. K. Veatch and G. T. DeTitta (2003) A deliberate approach to screening for initial crystallization conditions of biological macromolecules. *Journal of Structural Biology* 142(1): 170-179

Luft, J. R., J. Wolfley, I. Jurisica, J. Glasgow, S. Fortier and G. T. DeTitta (2001) Macromolecular crystallization in a high throughput laboratory- the search phase. *Journal of Crystal Growth* 232: 591-595

Mack, D., H. Rohde, S. Dobinsky, J. Riedewald, M. Nedelmann, J. K. Knobloch, H. A. Elsner and H. H. Feucht (2000) Identification of three essential regulatory gene loci governing expression of *Staphylococcus epidermidis* polysaccharide intercellular adhesin and biofilm formation. *Infection and Immunity* 68(7): 3799-3807

Maitra, S., C. Price and R. Ganguly (2002) *Cyp6a8* of *Drosophila melanogaster*: gene structure, and sequence and functional analysis of the upstream DNA. *Insect Biochemistry and Molecular Biology* 32(8): 859-870

Maitra, S., S. M. Dombrowski, L. C. Waters and R. Ganguly (1996) Three second chromosome-linked clustered *Cyp6* genes show differential constitutive and barbital-induced expression in DDT-resistant and susceptible strains of *Drosophila melanogaster*. *Gene* 180(1-2): 165-171

Malmqvist, M. and R. Karlsson (1997) Biomolecular interaction analysis: affinity biosensor technologies for functional analysis of proteins. *Current Opinion in Chemical Biology* 1(3): 378-383

Manavalan, P. and W. C. J. Johnson (1983) Sensitivity of circular dichroism to protein tertiary structure class. *Nature* 305: 831-832

Manukhov, I. V., G. E. Eroshnikov, M. Y. Vyssokikh and G. B. Zavilgelsky (1999) Folding and refolding of thermolabile and thermostable bacterial luciferases: the role of DnaKJ heat-shock proteins. *FEBS Letters* 448(2-3): 265-268

Mao, D., E. Wachter and B. A. Wallace (1982) Folding of the mitochondrial proton adenosine triphosphatase proteolipid channel in phospholipid vesicles. *Biochemistry* 21(20): 4960-4968

Markley, J. L. (1989) Two-dimensional nuclear magnetic resonance spectroscopy of proteins: An overview. Nuclear Magnetic Resonance Part A: Spectral Techniques and Dynamics. N. J. Oppenheimer and T. L. James. San Diego, Academic press. 176: 12-64

Marokhazi, J., G. Koczan, F. Hudecz, L. Graf, A. Fodor and I. Venekei (2004a) Enzymic characterization with progress curve analysis of a collagen peptidase from an enthomopathogenic bacterium, *Photorhabdus luminescens*. *Biochemical Journal* 379: 633-640

Marokhazi, J., K. Lengyel, S. Pekar, G. Felfoldi, A. Patthy, L. Graf, A. Fodor and I. Venekei (2004b) Comparison of proteolytic activities produced by entomopathogenic *Photorhabdus* bacteria: Strain- and phase-dependent heterogeneity in composition and activity of four enzymes. *Applied and Environmental Microbiology* 70(12): 7311-7320

Massey, R. C., A. Buckling and R. H. ffrench-Constant (2004) Interference competition and parasite virulence. *Proceedings of the Royal Society of London Series B* 271(1541): 785-788

McCart, C. (2006) Insecticide resistance in *Drosophila melanogaster* and *Ctenocephalides felis*. PhD thesis, Department of Biology and Biochemistry. University of Bath, Bath

McCart, C., A. Buckling and R. H. ffrench-Constant (2005) DDT resistance in flies carries no cost. *Current Biology* 15(15): R587-R589

McCrary, B. S., S. P. Edmondson and J. W. Shriver (1996) Hyperthermophile protein folding thermodynamics: differential scanning calorimetry and chemical denaturation of Sac7d. *Journal of Molecular Biology* 264(4): 784-805

Meiler, J. and D. Baker (2003) Coupled prediction of protein secondary and tertiary structure. *Proceedings of the National Academy of Sciences* 100(21): 12105-12110

Menozzi, P., M. Shi, A. Lougarre, Z. H. Tang and D. Fournier (2004) Mutations of acetylcholinesterase which confer insecticide resistance in *Drosophila melanogaster* populations. *BMC Evolutionary Biology* 4: 4

Meslet-Cladiere, L. M., A. Pimenta, E. Duchaud, I. B. Holland and M. A. Blight (2004) *In vivo* expression of the mannose-resistant fimbriae of *Photorhabdus temperata* K122 during insect infection. *Journal of Bacteriology* 186(3): 611-622

Mestres, J. (2005) Structure conservation in cytochromes P450. *Proteins* 58(3): 596-609

Miles, A. J., L. Whitmore and B. A. Wallace (2005) Spectral magnitude effects on the analyses of secondary structure from circular dichroism spectroscopic data. *Protein Science* 14(2): 368-374

Milstead, J. E. (1979) *Heterorhabditis bacteriophora* as a vector for introducing its associated bacterium into the hemocele of *Galleria mellonella* larvae. *Journal of Invertebrate Pathology* 33(3): 324-327

Miranker, A. D. and C. M. Dobson (1996) Collapse and cooperativity in protein folding. *Current Opinion in Structural Biology* 6(1): 31-42

Miyazaki, K. (2005) A hyperthermophilic laccase from *Thermus thermophilus* HB27. *Extremophiles* 9(6): 415-425

Moellenbeck, D. J., M. L. Peters, J. W. Bing, J. R. Rouse, L. S. Higgins, L. Sims, T. Nevshemal, L. Marshall, R. T. Ellis, P. G. Bystrak, B. A. Lang, J. L. Stewart, K. Kouba, V. Sondag, V. Gustafson, K. Nour, D. Xu, J. Swenson, J. Zhang, T. Czapla, G. Schwab, S. Jayne, B. A. Stockhoff, K. Narva, H. E. Schnepf, S. J. Stelman, C. Poutre, M. Koziel and N. Duck (2001) Insecticidal proteins from *Bacillus thuringiensis* protect corn from corn rootworms. *Nature Biotechnology* 19(7): 668-672

Monroe, K. R., S. P. Murphy, L. N. Kolonel and M. C. Pike (2007) Prospective study of grapefruit intake and risk of breast cancer in postmenopausal women: the Multiethnic Cohort Study. *British Journal of Cancer* 97(3): 440-445

Morgan, J. A. W., M. Sergeant, D. Ellis, M. Ousley and P. Jarrett (2001) Sequence analysis of insecticidal genes from *Xenorhabdus nematophilus* PMFI296. *Applied and Environmental Microbiology* 67(5): 2062-2069

Muralidhara, B. K., S. Negi, C. C. Chin, W. Braun and J. R. Halpert (2006) Conformational flexibility of mammalian cytochrome P450 2B4 in binding imidazole inhibitors with different ring chemistry and side chains. Solution thermodynamics and molecular modeling. *Journal of Biological Chemistry* 281(12): 8051-8061

Murray, T. S. and B. I. Kazmierczak (2006) FlhF is required for swimming and swarming in *Pseudomonas aeruginosa*. *Journal of Bacteriology* 188(19): 6995-7004

Muthusamy, R., M. M. Gromiha and P. K. Ponnuswamy (2000) On the thermal unfolding character of globular proteins. *Journal of Protein Chemistry* 19(1): 1-8

Nikou, D., H. Ranson and J. Hemingway (2003) An adult-specific CYP6 P450 gene is overexpressed in a pyrethroid-resistant strain of the malaria vector, *Anopheles gambiae*. *Gene* 318: 91-102

Nilsson, M. R. (2004) Techniques to study amyloid fibril formation *in vitro*. *Methods* 34(1): 151-160

Nelson, D. R. (2007) Cytochrome P450 family count. Available online at http://drnelson.utmem.edu/p450stats.2007.htm

Nelson, D., L. Koymans, i. T. Kamatak, J. Stegeman, R. Feyereisen, D. Waxman, M. Waterman, O. Gotoh, M. Coon, R. Estabrook, I. Gunsalus and N. DW (1996) P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics* 6(1): 1-42

Neu, T. R. (1996) Significance of bacterial surface-active compounds in interaction of bacteria with interfaces. *Microbiological Reviews* 60(1): 151-166

Nishimoto, M., J. E. Clark and B. S. Masters (1993) Cytochrome P450 4A4: expression in *Escherichia coli*, purification, and characterization of catalytic properties. *Biochemistry* 32(34): 8863-8870

Nominé, Y., T. Ristriani, C. Laurent, J. F. Lefevre, E. Weiss and G. Trave (2001) A strategy for optimizing the monodispersity of fusion proteins: application to purification of recombinant HPV E6 oncoprotein. *Protein Engineering* 14(4): 297-305

Nougayrede, J. P., F. Taieb, J. De Rycke and E. Oswald (2005) Cyclomodulins: bacterial effectors that modulate the eukaryotic cell cycle. *Trends in Microbiology* 13(3): 103-110

O'Neill, K. H., D. M. Roche, D. J. Clarke and B. C. A. Dowds (2002) The *ner* gene of *Photorhabdus*: Effects on primary-form-specific phenotypes and outer membrane protein composition. *Journal of Bacteriology* 184(14): 3096-3105

O'Quinn, A. L., E. M. Wiegand and J. A. Jeddeloh (2001) *Burkholderia pseudomallei* kills the nematode *Caenorhabditis elegans* using an endotoxin-mediated paralysis. *Cellular Microbiology* 3(6): 381-393

O'Toole, G. A. and R. Kolter (1998) Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. *Molecular Microbiology* 30(2): 295-304

Otwinowski, Z. (1993) SCALEPACK. Yale University, New Haven, Connecticut

Overhage, J., S. Lewenza, A. K. Marr and R. E. W. Hancock (2007) Identification of genes involved in swarming motility using a *Pseudomonas aeruginosa* PAO1 mini-Tn5-lux mutant library. *Journal of Bacteriology* 189(5): 2164-2169

Pamp, S. J. and T. Tolker-Nielsen (2007) Multiple roles of biosurfactants in structural biofilm development by *Pseudomonas aeruginosa*. *Journal of Bacteriology* 189(6): 2531-2539

Pan, K., M. Baldwin, J. Nguyen, M. Gasset, A. Serban, D. Groth, I. Mehlhorn, Z. Huang, R. Fletterick, F. Cohen and S. Prusiner (1993) Conversion of alpha-helices beta-sheets features in the formation of the scrapie prion proteins. *Proceedings of the National Academy of Sciences* 90(23): 10962-10966

Park, Y. I. and A. K. Raina (2004) Juvenile hormone III titers and regulation of soldier caste in *coptotermes formosanus* (Isoptera: Rhinotermitidae). *Journal of Insect Physiology* 50(6): 561-566

Parkhill, J., B. W. Wren, N. R. Thomson, R. W. Titball, M. T. G. Holden, M. B. Prentice, M. Sebaihia, K. D. James, C. Churcher, K. L. Mungall, S. Baker, D. Basham, S. D. Bentley, K. Brooks, A. M. Cerdeno-Tarraga, T. Chillingworth, A. Cronin, R. M. Davies, P. Davis, G. Dougan, T. Feltwell, N. Hamlin, S. Holroyd, K. Jagels, A. V. Karlyshev, S. Leather, S. Moule, P. C. F. Oyston, M. Quail, K. Rutherford, M. Simmonds, J. Skelton, K. Stevens, S. Whitehead and B. G. Barrell (2001) Genome sequence of *Yersinia pestis*, the causative agent of plague. *Nature* 413(6855): 523-527

Parsek, M. R. and E. P. Greenberg (2000) Acyl-homoserine lactone quorum sensing in Gram-negative bacteria: A signaling mechanism involved in associations with higher organisms. *Proceedings of the National Academy of Sciences* 97(16): 8789-8793

Pasmore, M. and J. W. Costerton (2003) Biofilms, bacterial signaling, and their ties to marine biology. *Journal of Industrial Microbiology and Biotechnology* 30(7): 407-413

Patel, J. D., M. Ebert, R. Ward and J. M. Anderson (2007) *S. epidermidis* biofilm formation: effects of biomaterial surface chemistry and serum proteins. *Journal of Biomedical Materials Research, Part A* 80(3): 742-751

Pawar, A. P., K. F. Dubay, J. Zurdo, F. Chiti, M. Vendruscolo and C. M. Dobson (2005) Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases. *Journal of Molecular Biology* 350(2): 379-392

Petersen, R. A., A. R. Zangerl, M. R. Berenbaum and M. A. Schuler (2001) Expression of CYP6B1 and CYP6B3 cytochrome P450 monooxygenases and furanocoumarin metabolism in different tissues of *Papilio polyxenes* (Lepidoptera: Papilionidae). *Insect Biochemistry and Molecular Biology* 31(6-7): 679-690

Petryk, A., J. T. Warren, G. Marques, M. P. Jarcho, L. I. Gilbert, J. Kahler, J. P. Parvy, Y. Li, C. Dauphin-Villemant and M. B. O'Connor (2003) *Shade* is the *Drosophila* P450 enzyme that mediates the hydroxylation of ecdysone to the steroid insect molting hormone 20-hydroxyecdysone. *Proceedings of the National Academy of Sciences* 100(24): 13773-13778

Piszczek, G., S. D'Auria, M. Staiano, M. Rossi and A. Ginsburg (2004) Conformational stability and domain coupling in D-glucose/D-galactose-binding protein from *Escherichia coli*. *Biochemical Journal* 381(1): 97-103

Ploux, L., S. Beckendorff, M. Nardin and S. Neunlist (2007) Quantitative and morphological analysis of biofilm formation on self-assembled monolayers. *Colloids and Surfaces B: Biointerfaces* 57(2): 174-181

Poulos, T. L. (2003) Cytochrome P450 flexibility. *Proceedings of the National Academy of Sciences* 100(23): 13121-13122

Poulos, T. L., B. C. Finzel, I. C. Gunsalus, G. C. Wagner and J. Kraut (1985) The 2.6-A crystal structure of *Pseudomonas putida* cytochrome P-450. *Journal of Biological Chemistry* 260(30): 16122-16130

Pratt, L. A. and R. Kolter (1998) Genetic analysis of *Escherichia coli* biofilm formation: roles of flagella, motility, chemotaxis and type I pili. *Molecular Microbiology* 30(2): 285-293

Pusey, P. N. and R. J. A. Tough (1985) Chapter 4, Particle interactions. Dynamic Light Scattering. R. Pecora. New York, Plenum Press: 85-179

Pyke, F. M., M. R. Bogwitz, T. Perry, A. Monk, P. Batterham and J. A. McKenzie (2004) The genetic basis of resistance to diazinon in natural populations of *Drosophila melanogaster*. *Genetica* 121(1): 13-24

Querol, E., J. A. Perez-Pons and A. Mozo-Villarias (1996) Analysis of protein conformational characteristics related to thermostability. *Protein Engineering* 9(3): 265-271

Ramia, S., E. Neter and D. Brenner (1982) Production of enterobacterial common antigen as an aid to classification of newly identified species of the families *Enterbacteriaceae* and *Vibrioneceae*. *International Journal of Systematic Bacteriology* 32: 395-398

Ranasinghe, C., B. Campbell and A. A. Hobbs (1998) Over-expression of cytochrome P450 CYP6B7 mRNA and pyrethroid resistance in Australian populations of *Helicoverpa armigera* (Hübner). *Pesticide Science* 54(3): 195-202

Religa, T. L., J. S. Markson, U. Mayor, S. M. V. Freund and A. R. Fersht (2005) Solution structure of a protein denatured state and folding intermediate. 437(7061): 1053-1056

Rewitz, K. F., R. Rybczynski, J. T. Warren and L. I. Gilbert (2006) The Halloween genes code for cytochrome P450 enzymes mediating synthesis of the insect moulting hormone. *Biochemical Society Transactions* 34(6): 1256-1260

Robic, S., M. Guzman-Casado, J. M. Sanchez-Ruiz and S. Marqusee (2003) Role of residual structure in the unfolded state of a thermophilic protein. *Proceedings of the National Academy of Sciences* 100(20): 11345-11349

Rochu, D., N. Beaufet, F. Renault, N. Viguie and P. Masson (2002) The wild type bacterial Co(2+)/Co(2+)-phosphotriesterase shows a middle-range thermostability. *Biochimica et Biophysica Acta* 1594(2): 207-218

Rohl, C. A., J. M. Scholtz, E. J. York, J. M. Stewart and R. L. Baldwin (1992) Kinetics of amide proton exchange in helical peptides of varying chain lengths. Interpretation by the Lifson-Roig equation. *Biochemistry* 31(5): 1263-1269

Ron, E. Z. and E. Rosenberg (2001) Natural roles of biosurfactants. *Environmental Microbiology* 3(4): 229-236

Rost, B., G. Yachdav and L. Jinfeng (2004) The PredictProtein server. *Nucleic Acids Research* 32(Web server issue): W321-W326

Rousseau, F., L. Serrano and J. W. Schymkowitz (2006a) How evolutionary pressure against protein aggregation shaped chaperone specificity. *Journal of Molecular Biology* 355(5): 1037-1047

Rousseau, F., J. Schymkowitz and L. Serrano (2006b) Protein aggregation and amyloidosis: confusion of the kinds? *Current Opinion in Structural Biology* 16(1): 118-126

Ruiz-Arribas, A., G. G. Zhadan, V. P. Kutyshenko, R. I. Santamaria, M. Cortijo, E. Villar, J. M. Fernandez-Abalos, J. J. Calvete and V. L. Shnyrov (1998) Thermodynamic stability of two variants of xylanase (Xys1) from *Streptomyces halstedii* JM8. *European Journal of Biochemistry* 253(2): 462-468

Sambrook, J., E. F. Fritsch and T. Maniatis (1989) Molecular cloning. Cold Spring Harbor, Cold Spring Harbor Press

Savchenko, T., P. Whiting, A. Germade and L. Dinan (2000) Ecdysteroid agonist and antagonist activities in species of the Solanaceae. *Biochemical Systematics and Ecology* 28(5): 403-419

Schlenke, T. A. and D. J. Begun (2004) Strong selective sweep associated with a transposon insertion in *Drosophila simulans*. *Proceedings of the National Academy of Sciences* 101(6): 1626-1631

Schmitz, R. A. (1997) NifL of *Klebsiella pneumoniae* carries an N-terminally bound FAD cofactor, which is not directly required for the inhibitory function of NifL. *FEMS Microbiology Letters* 157(2): 313-318

Schnepf, H. E., S. Lee, J. Dojillo, P. Burmeister, K. Fencil, L. Morera, L. Nygaard, K. E. Narva and J. D. Wolt (2005) Characterization of Cry34/Cry35 binary insecticidal proteins from diverse *Bacillus thuringiensis* strain collections. *Applied and Environmental Microbiology* 71(4): 1765-1774

Schoch, G. A., J. K. Yano, M. R. Wester, K. J. Griffin, C. D. Stout and E. F. Johnson (2004) Structure of human microsomal cytochrome P450 2C8. Evidence for a peripheral fatty acid binding site. *Journal of Biological Chemistry* 279(10): 9497-9503

Schooling, S. R. and T. J. Beveridge (2006) Membrane vesicles: an overlooked component of the matrices of biofilms. *Journal of Bacteriology* 188(16): 5945-5957

Schwartz, R., C. S. Ting and J. King (2001) Whole proteome pI values correlate with subcellular localizations of proteins for organisms within the three domains of life. *Genome Research* 11(5): 703-709

Scott, E. E. and J. R. Halpert (2005) Structures of cytochrome P450 3A4. *Trends in Biochemical Sciences* 30(1): 5-7

Scott, E. E., Y. A. He, M. R. Wester, M. A. White, C. C. Chin, J. R. Halpert, E. F. Johnson and C. D. Stout (2003) An open conformation of mammalian cytochrome P450 2B4 at 1.6-A resolution. *Proceedings of the National Academy of Sciences* 100(23): 13196-13201

Scott, E. E., M. Spatzenegger and J. R. Halpert (2001) A truncation of 2B subfamily cytochromes P450 yields increased expression levels, increased solubility, and decreased aggregation while retaining function. *Archives of Biochemistry and Biophysics* 395(1): 57-68

Shallom, D., G. Golan, G. Shoham and Y. Shoham (2004) Effect of dimer dissociation on activity and thermostability of the alpha-glucuronidase from *Geobacillus stearothermophilus*: dissecting the different oligomeric forms of family 67 glycoside hydrolases. *Journal of Bacteriology* 186(20): 6928-6937

Sharma, S., N. Waterfield, D. Bowen, T. Rocheleau, L. Holland, R. James and R. ffrench-Constant (2002) The lumicins: novel bacteriocins from *Photorhabdus luminescens* with similarity to the uropathogenic-specific protein (USP) from uropathogenic *Escherichia coli*. *FEMS Microbiology Letters* 214(2): 241-249

Sheffield, P., S. Garrard and Z. Derewenda (1999) Overcoming expression and purification problems of RhoGDI using a family of "parallel" expression vectors. *Protein Expression and Purification* 15(1): 34-39

Shen, B., H.-Q. Dong, H.-S. Tian, L. Ma, X.-L. Li, G.-L. Wu and C.-L. Zhu (2003) Cytochrome P450 genes expressed in the deltamethrin-susceptible and -resistant strains of *Culex pipiens pallens*. *Pesticide Biochemistry and Physiology* 75(1-2): 19-26

Shi, M. A., A. Lougarre, C. Alies, I. Fremaux, Z. H. Tang, J. Stojan and D. Fournier (2004) Acetylcholinesterase alterations reveal the fitness cost of mutations conferring insecticide resistance. *BMC Evolutionary Biology* 4: 5

Shortle, D. (1996) The denatured state (the other half of the folding equation) and its role in protein stability. *FASEB Journal* 10(1): 27-34

Sicard, M., S. Hering, R. Schulte, S. Gaudriault and H. Schulenburg (2007) The effect of *Photorhabdus luminescens* (*Enterobacteriaceae*) on the survival, development, reproduction and behaviour of *Caenorhabditis elegans* (Nematoda: *Rhabditidae*). *Environmental Microbiology* 9(1): 12-25

Silva, C. P., N. R. Waterfield, P. J. Daborn, P. Dean, T. Chilver, C. P. Y. Au, S. Sharma, U. Potter, S. E. Reynolds and R. H. ffrench-Constant (2002) Bacterial infection of a model insect: *Photorhabdus luminescens* and *Manduca sexta*. *Cellular Microbiology* 4(6): 329-339

Sipe, J. D. and A. S. Cohen (2000) Review: history of the amyloid fibril. *Journal of Structural Biology* 130(2-3): 88-98

Smigielski, A. J. and R. J. Akhurst (1994) Megaplasmids in *Xenorhabdus* and *Photorhabdus* spp, bacterial symbionts of entomopathogenic nematodes (families *Steinernematidae* and *Heterorhabditidae*). *Journal of Invertebrate Pathology* 64(3): 214-220

Solano, C., B. Garcia, J. Valle, C. Berasain, J. M. Ghigo, C. Gamazo and I. Lasa (2002) Genetic analysis of *Salmonella enteritidis* biofilm formation: critical role of cellulose. *Molecular Microbiology* 43(3): 793-808

Souza, D. H. F., H. S. Selistre-de-Araujo and R. C. Garratt (2000) Determination of the three-dimensional structure of toxins by protein crystallography. *Toxicon* 38(10): 1307-1353

Spraul, M., A. S. Freund, R. E. Nast, R. S. Withers, W. E. Maas and O. Corcoran (2003) Advancing NMR sensitivity for LC-NMR-MS using a cryoflow probe: application to the analysis of acetaminophen metabolites in urine. *Analytical Chemistry* 75(6): 1536-1541

Sreerama, N., S. Y. Venyaminov and R. W. Woody (1999) Estimation of the number of alpha-helical and beta-strand segments in proteins using circular dichroism spectroscopy. *Protein Science* 8: 370-380

Srisailam, S., T. K. Kumar, D. Rajalingam, K. M. Kathir, H. S. Sheu, F. J. Jan, P. C. Chao and C. Yu (2003) Amyloid-like fibril formation in an all beta-barrel protein. Partially structured intermediate state(s) is a precursor for fibril formation. *Journal of Biological Chemistry* 278(20): 17701-17709

Stanger, H. E., F. A. Syud, J. F. Espinosa, I. Giriat, T. Muir and S. H. Gellman (2001) Length-dependent stability and strand length limits in antiparallel beta -sheet secondary structure. *Proceedings of the National Academy of Sciences* 98(21): 12015-12020

Stout, C. D. (2004) Cytochrome P450 conformational diversity. *Structure* 12(11): 1921-1922

Sturtevant, J. M. (1977) Heat capacity and entropy changes in processes involving proteins. *Proceedings of the National Academy of Sciences* 74(6): 2236-2240

Sutherland, I. (2001) Biofilm exopolysaccharides: a strong and sticky framework. *Microbiology* 147(1): 3-9

Szabo, A., L. Stolz and R. Granzow (1995) Surface plasmon resonance and its use in biomolecular interaction analysis (Bia). *Current Opinion in Structural Biology* 5(5): 699-705

Szállás, E., C. Koch, A. Fodor, J. Burghardt, O. Buss, A. Szentirmai, K. H. Nealson and E. Stackebrandt (1997) Phylogenetic evidence for the taxonomic heterogeneity of *Photorhabdus luminescens*. *International Journal of Systematic Bacteriology* 47(2): 402-407

Szittner, R. and E. Meighen (1990) Nucleotide-sequence, expression, and properties of luciferase coded by lux genes from a terrestrial bacterium. *Journal of Biological Chemistry* 265(27): 16581-16587

Takeuchi, K., E. Ng, T. J. Malia and G. Wagner (2007) 1-^{13}C amino acid selective labeling in a $^{2}H^{15}N$ background for NMR studies of large proteins. *Journal of Biolmolecular NMR* 38((1)): 89-98

Tan, L. and C. Darby (2004) A movable surface: Formation of *Yersinia* sp biofilms on motile *Caenorhabditis elegans*. *Journal of Bacteriology* 186(15): 5087-5092

Tegoulia, V. A. and S. L. Cooper (2002) *Staphylococcus aureus* adhesion to self-assembled monolayers: effect of surface chemistry and fibrinogen presence. *Colloids and Surfaces B-Biointerfaces* 24(3-4): 217-228

Thomas, G. M. and G. O. Poinar Jr (1979) *Xenorhabdus* gen. nov., a genus of entomopathogenic nematophilic bacteria of the family *Enterobacteriaceae*. *International Journal of Systematic Bacteriology* 29: 352-360

Thomas, T. and R. Cavicchioli (2000) Effect of temperature on stability and activity of elongation factor 2 proteins from Antarctic and thermophilic methanogens. *Journal of Bacteriology* 182(5): 1328-1332

Tijet, N., C. Helvig and R. Feyereisen (2001) The cytochrome P450 gene superfamily in *Drosophila melanogaster*: Annotation, intron-exon organization and phylogeny. *Gene* 262(1-2): 189-198

Torreri, P., M. Ceccarini, P. Macioce and T. Petrucci (2005) Biomolecular interactions by surface plasmon resonance technology. *Annali dell'Istituto Superiore di Sanita* 41(4): 437-441

Tounsi, S., M. Blight, S. Jaoua and A. de Lima Pimenta (2006) From insects to human hosts: Identification of major genomic differences between entomopathogenic strains of *Photorhabdus* and the emerging human pathogen *Photorhabdus asymbiotica. International Journal of Medical Microbiology* 296(8): 521-530

Turlin, E., G. Pascal, J. C. Rousselle, P. Lenormand, S. Ngo, A. Danchin and S. Derzelle (2006) Proteome analysis of the phenotypic variation process in *Photorhabdus luminescens. Proteomics* 6(9): 2705-2725

van Hamme, J. D., A. Singh and O. P. Ward (2006) Physiological aspects. Part 1 in a series of papers devoted to surfactants in microbiology and biotechnology. *Biotechnology Advances* 24(6): 604-620

van Holde, K. E., W. C. Johnson and P. S. Ho (1998) Chapter 10, Linear and circular dichroism. Principles of physical biochemistry. Upper Saddle River, NJ, Prentice Hall: 418-451

van Nuland, N. A. J., F. Chiti, N. Taddei, G. Raugei, G. Ramponi and C. M. Dobson (1998) Slow folding of muscle acylphosphatase in the absence of intermediates. *Journal of Molecular Biology* 283(4): 883-891

Vojtěchovský, J., K. Chu, J. Berendzen, R. M. Sweet and I. Schlichting (1999) Crystal structures of myoglobin-ligand complexes at near-atomic resolution. *Biophysical Journal* 77(4): 2153-2174

Wang, Y. and R. Gaugler (1998) Host and penetration site location by entomopathogenic nematodes against Japanese beetle larvae. *Journal of Invertebrate Pathology* 72(3): 313-318

Ware, G. W. (2004) An introduction to insecticides. Radcliffe's IPM world textbook. E. B. Radcliffe and W. D. Hutchison. St. Paul, MN, Available online at http//www.impworld.umn.edu/chapters/ware.htm

Waterfield, N., M. Hares, G. Yang, A. Dowling and R. ffrench-Constant (2005) Potentiation and cellular phenotypes of the insecticidal toxin complexes of *Photorhabdus bacteria. Cellular Microbiology* 7(3): 373-382

Waterfield, N. R., D. J. Bowen, J. D. Fetherston, R. D. Perry and R. H. ffrench-Constant (2001) The tc genes of *Photorhabdus*: a growing family. *Trends in Microbiology* 9(4): 185-191

Waterfield, N. R., P. J. Daborn, A. J. Dowling, G. W. Yang, M. Hares and R. H. ffrench-Constant (2003) The insecticidal toxin makes caterpillars floppy 2 (Mcf2) shows similarity to HrmA, an avirulence protein from a plant pathogen. *Fems Microbiology Letters* 229(2): 265-270

Waterfield, N. R., P. J. Daborn and R. H. ffrench-Constant (2002) Genomic islands in *Photorhabdus*. *Trends in Microbiology* 10(12): 541-545

Weissfeld, A. S., R. J. Halliday, D. E. Simmons, E. A. Trevino, P. H. Vance, C. M. O'Hara, E. G. Sowers, R. Kern, R. D. Koy, K. Hodde, M. Bing, C. Lo, J. Gerrard, R. Vohra and J. Harper (2005) *Photorhabdus asymbiotica*, a pathogen emerging on two continents that proves that there is no substitute for a well-trained clinical microbiologist. *Journal Clinical Microbiology* 43(8): 4152-4155

Werck-Reichhart, D. and R. Feyereisen (2000) Cytochromes P450: a success story. *Genome Biology* 1(6): REVIEWS3003

Wester, M. R., E. F. Johnson, C. Marques-Soares, S. Dijols, P. M. Dansette, D. Mansuy and C. D. Stout (2003) Structure of mammalian cytochrome P450 2C5 complexed with diclofenac at 2.1 A resolution: evidence for an induced fit model of substrate binding. *Biochemistry* 42(31): 9335-9345

Whitford, D. (2005a) Chapter 3, The three-dimensional structure of proteins. Protein structure and function. D. Whitford. Chichester, Wiley and Sons: 39-84

Whitford, D. (2005b) Chapter 11, Protein folding *in vivo* and *in vitro*. Protein structure and function. D. Whitford. Chichester, Wiley and Sons: 395-438

Whitmore, L. and B. A. Wallace (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Research* 32(Web Server issue): W668-673

Wider, G. (2005) NMR techniques used with very large biological macromolecules in solution. Nuclear magnetic resonance of biological macromolecules, part C. T. L. James. Can Diego, California, Elsevier Academic Press. 394: 382-398

Wiencek, J. M. (1999) New Strategies for Protein Crystal Growth. *Annual Review of Biomedical Engineering* 1(1): 505-534

Williams, J. S., M. Thomas and D. J. Clarke (2005) The gene stlA encodes a phenylalanine ammonia-lyase that is involved in the production of a stilbene antibiotic in *Photorhabdus luminescens* TT01. *Microbiology* 151(8): 2543-2550

Williams, P. A., J. Cosme, A. Ward, H. C. Angove, D. Matak Vinkovic and H. Jhoti (2003) Crystal structure of human cytochrome P450 2C9 with bound warfarin. *Nature* 424(6947): 464-468

Williams, P. A., J. Cosme, D. M. Vinkovic, A. Ward, H. C. Angove, P. J. Day, C. Vonrhein, I. J. Tickle and H. Jhoti (2004) Crystal structures of human cytochrome P450 3A4 bound to metyrapone and progesterone. *Science* 305(5684): 683-686

Williamson, M. S., D. Martinez-Torres, C. A. Hick and A. L. Devonshire (1996) Identification of mutations in the housefly *para*-type sodium channel gene associated with knockdown resistance (*kdr*) to pyrethroid insecticides. *Molecular and General Genetics* 27(1-2): 51-60

Willott, E., T. Trenczek, L. W. Thrower and M. R. Kanost (1994) Immunochemical identification of insect hemocyte populations: monoclonal antibodies distinguish four major hemocyte types in *Manduca sexta*. *European Journal of Cell Biology* 65(2): 417-423

Willoughby, L., H. Chung, C. Lumb, C. Robin, P. Batterham and P. J. Daborn (2006) A comparison of *Drosophila melanogaster* detoxification gene induction responses for six insecticides, caffeine and phenobarbital. *Insect Biochemistry and Molecular Biology* 36(12): 934-942

Wishart, D. S. and D. A. Case (2001) Use of chemical shifts in macromolecular structure determination. Nuclear Magnetic Resonance of Biological Macromolecules. T. L. James, V. Dotsch and U. Schmitz. San Diego, Academic press. 338: 3-34

Yang, G., A. J. Dowling, U. Gerike, R. H. ffrench-Constant and N. R. Waterfield (2006) *Photorhabdus* virulence cassettes confer injectable insecticidal activity against the wax moth. *Journal of Bacteriology* 188(6): 2254-2261

Yano, J. K., M. R. Wester, G. A. Schoch, K. J. Griffin, C. D. Stout and E. F. Johnson (2004) The structure of human microsomal cytochrome P450 3A4 determined by X-ray crystallography to 2.05Å resolution. *Journal of Biological Chemistry* 279(37): 38091-38094

Yee, A. A., A. Savchenko, A. Ignachenko, J. Lukin, X. Xu, T. Skarina, E. Evdokimova, C. S. Liu, A. Semesi, V. Guido, A. M. Edwards and C. H. Arrowsmith (2005) NMR and X-ray crystallography, complementary tools in structural proteomics of small proteins. *Journal of the American Chemical Society* 127(47): 16512-16517

Zerbe, K., O. Pylypenko, F. Vitali, W. Zhang, S. Rouset, M. Heck, J. W. Vrijbloed, D. Bischoff, B. Bister, R. D. Sussmuth, S. Pelzer, W. Wohlleben, J. A. Robinson and I. Schlichting (2002) Crystal structure of OxyB, a cytochrome P450 implicated in an oxidative phenol coupling reaction during vancomycin biosynthesis. *Journal of Biological Chemistry* 277(49): 47476-47485

Zhu, X., P. Wentworth, Jr., R. A. Kyle, R. A. Lerner and I. A. Wilson (2006) Cofactor-containing antibodies: Crystal structure of the original yellow antibody. *Proceedings of the National Academy of Sciences* 103(10): 3581-3585

Zimm, B. H., P. Doty and K. Iso (1959) Determination of the parameters for helix formation in poly-gamma-benzyl-L-glutamate. *Proceedings of the National Academy of Sciences of the USA* 45(11): 1601-1607

Zogaj, X., M. Nimtz, M. Rohde, W. Bokranz and U. Römling (2001) The multicellular morphotypes of *Salmonella typhimurium* and *Escherichia coli* produce cellulose as the second component of the extracellular matrix. *Molecular Microbiology* 39(6): 1452-1463